Dupont L, Pallot M, Morel L (2016) Exploring the Appropriateness of Different Immersive Environments in the Context of an Innovation Process for Smart Cities. In 22nd ICE/IEEE International Technology Management Conference. Trondheim, Norway.
Dykes J, MacEachren AM, Kraak MJ (2005a) Exploring Geovisualization. In Exploring Geovisualization, edited by Jason Dykes, Alan M. MacEachren, and Menno-Jan Kraak, 1–19. Amsterdam: Pergamon Press.
Dykes J, MacEachren AM, Kraak MJ (2005b) Advancing Geovisualization. In Exploring Geovisualization, edited by Jason Dykes, Alan M. MacEachren, and Menno-Jan Kraak, 691–703. Amsterdam: Pergamon Press. https://doi.org/10.1016/b978-008044531-1/50454-1.
Fairbairn D (2006) Measuring Map Complexity. The Cartographic Journal 43: 224–238. https://doi.org/10.1179/000870406x169883.
Fisher P, Unwin D (2001) Virtual Reality in Geography. London: CRC Press.
Fuhrmann S, Paula AR, Edsall RM et al (2005) Making Useful and Useable Geovisualization: Design and Evaluation Issues. In Exploring Geovisualization, edited by Jason Dykes, Alan M. MacEachren, and Menno-Jan Kraak, 553–566. Pergamon Press. https://doi.org/10.1016/b978-008044531-1/50446-2.
Gabbard JL, Hix D, Swan EJ (1999) User-Centered Design and Evaluation of Virtual Environments. IEEE Computer Graphics and Applications 19: 51–59. https://doi.org/10.1109/38.799740.
Gandomi A, Haider M (2015) Beyond the Hype: Big Data Concepts, Methods, and Analytics. International Journal of Information Management 35: 137–144. https://doi.org/10.1016/j.ijinfomgt.2014.10.007.
Garlandini S, Fabrikant SI (2009) Evaluating the Effectiveness and Efficiency of Visual Variables for Geographic Information Visualization. In COSIT 2009, Lecture Notes in Computer Science, 5756:195–211. Berlin Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-642-03832-7_12.
Geertman S, Allan A, Pettit C et al (2017) Introduction to Planning Support Science for Smarter Urban Futures. In Planning Support Science for Smarter Urban Futures, edited by Stan Geertman, Andrew Allan, Chris Pettit, and Stillwell, John, 1–19. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-57819-4.
Goodchild MF (2007) Citizens as Sensors: The World of Volunteered Geography. GeoJournal 69: 211–221. https://doi.org/10.1007/s10708-007-9111-y.
Goodchild MF (2009) Virtual Geographic Environments as Collective Constructions. In Virtual Geographic Environments, edited by Hui Lin and Michael Batty, 15–24. Beijing: Science Press.
Goodchild MF, Guo HD, Annoni A et al (2012) Next-Generation Digital Earth. Proceedings of the National Academy of Sciences 109: 11088–11094. https://doi.org/10.1073/pnas.1202383109.
Gore A (1998) The Digital Earth: Understanding Our Planet in the 21st Century. Australian Surveyor 43 (2): 89–91. https://doi.org/10.1080/00050326.1998.10441850.
Gotow JB, Zienkiewicz K, White J et al (2010) Mobile Wireless Middleware, Operating Systems, and Applications, Third International Conference, Mobilware 2010, Chicago, IL, USA, June 30–July 2, 2010. Revised Selected Papers, 129–143. https://doi.org/10.1007/978-3-642-17758-3_10.
Gray S, O'Brien O, Hügel S (2016) Collecting and Visualizing Real-Time Urban Data through City Dashboards. Built Environment 42: 498–509. https://doi.org/10.2148/benv.42.3.498.
Griffin AL (2004) Understanding How Scientists Use Data-Display Devices for Interactive Visual Computing with Geographical Models. PhD Thesis, University Park, PA: Department of Geography, The Pennsylvania State University.
Griffin AL, Fabrikant SI (2012) More Maps, More Users, More Devices Means More Cartographic Challenges. The Cartographic Journal 49: 298–301. https://doi.org/10.1179/0008704112z.00000000049.
Griffin AL, White T, Fish C et al (2017) Designing across Map Use Contexts: A Research Agenda. International Journal of Cartography 3: 1–25. https://doi.org/10.1080/23729333.2017.1315988.
Grossner KE, Goodchild MF, Clarke KC (2008) Defining a Digital Earth System. Transactions in GIS 12: 145–160. https://doi.org/10.1111/j.1467-9671.2008.01090.x.

Hägerstrand T (1970) What about People in Regional Science? Papers in Regional Science 24: 7–24. https://doi.org/10.1111/j.1435-5597.1970.tb01464.x.
Harrower M, Brewer CA (2003) ColorBrewer.Org: An Online Tool for Selecting Colour Schemes for Maps. The Cartographic Journal 40: 27–37.
Hegarty M (2011) The Cognitive Science of Visual-Spatial Displays: Implications for Design. Topics in Cognitive Science 3: 446–474. https://doi.org/10.1111/j.1756-8765.2011.01150.x.
Hegarty M, Waller DA (2005) Individual Differences in Spatial Abilities. In The Cambridge Handbook of Visuospatial Thinking, edited by Priti Shah and Akira Miyake, 121–169. Cambridge, UK: Cambridge University Press.
Heim M (1998) Virtual Realism. Oxford: Oxford University Press.
Helbig C, Bauer HS, Rink K et al (2014) Concept and Workflow for 3D Visualization of Atmospheric Data in a Virtual Reality Environment for Analytical Approaches. Environmental Earth Sciences 72: 3767–3780. https://doi.org/10.1007/s12665-014-3136-6.
Heuer RJ (1999) Psychology of Intelligence Analysis. Langley, VA: Central Intelligence Agency.
Hoarau C, Christophe S (2017) Cartographic Continuum Rendering Based on Color and Texture Interpolation to Enhance Photo-Realism Perception. ISPRS Journal of Photogrammetry and Remote Sensing 127: 27–38. https://doi.org/10.1016/j.isprsjprs.2016.09.012.
Holten D, van Wijk JJ (2009) Force-Directed Edge Bundling for Graph Visualization. Computer Graphics Forum 28 (3): 983–990. https://doi.org/10.1111/j.1467-8659.2009.01450.x.
Huang HS, Schmidt M, Gartner G (2012) Spatial Knowledge Acquisition with Mobile Maps, Augmented Reality and Voice in the Context of GPS-Based Pedestrian Navigation: Results from a Field Test. Cartography and Geographic Information Science 39: 107–116. https://doi.org/10.1559/15230406392107.
Huang L, Gong JH, Li WH et al (2018) Social Force Model-Based Group Behavior Simulation in Virtual Geographic Environments. ISPRS International Journal of Geo-Information 7: article 79. https://doi.org/10.3390/ijgi7020079.
Hullman J (2016) Why Evaluating Uncertainty Visualization Is Error Prone. In BELIV'2016, 143–151. Baltimore, MD. https://doi.org/10.1145/2993901.2993919.
Hurter C (2015) Image-Based Visualisation: Interactive Multidimensional Data Exploration. London: Morgan & Claypool Publishers. https://doi.org/10.2200/s00688ed1v01y201512vis006.
Hurter C, Puechmorel S, Nicol F et al (2018) Functional Decomposition for Bundled Simplification of Trail Sets. IEEE Transactions on Visualization and Computer Graphics 24: 500–510. https://doi.org/10.1109/tvcg.2017.2744338.
Jacquinod F, Pedrinis F, Edert J et al (2016) Automated Production of Interactive 3D Temporal Geovisualizations so as to Enhance Flood Risk Awareness. In UDMV '16 Proceedings of the Eurographics Workshop on Urban Data Modelling and Visualisation, 71–77. Liège, Belgium: The Eurographics Association.
Jenny B, Kelso NV (2007) Color Design for the Color Vision Impaired. Cartographic Perspectives 58: 61–67. https://doi.org/10.14714/cp58.270.
Jenny B, Stephen DM, Muehlenhaus I et al (2017) Force-Directed Layout of Origin-Destination Flow Maps. International Journal of Geographical Information Science 31 (8): 1521–1540. https://doi.org/10.1080/13658816.2017.1307378.
Jenny B, Stephen DM, Muehlenhaus I et al (2018) Design Principles for Origin-Destination Flow Maps. Cartography and Geographic Information Science 45: 62–75. https://doi.org/10.1080/15230406.2016.1262280.
Jenny H, Jenny B, Cron J (2012) Exploring Transition Textures for Pseudo-Natural Maps. In GI_Forum 2012: Geovisualization, Society and Learning, edited by T. Jekel, A. Car, Josef Strobl, and G. Griesenebe r, 130–139. Berlin: Wichmann.
Jerald J (2015) The VR Book. New York: ACM & Morgan & Claypool Publishers. https://doi.org/10.1145/2792790.
Jia FL, You X, Tian JP et al (2015) Formal Language for the Virtual Geographic Environment. Environmental Earth Sciences 74: 6981–7002. https://doi.org/10.1007/s12665-015-4756-1.

Just MA, Carpenter PA (1976) The Role of Eye-Fixation Research in Cognitive Psychology. Behavior Research Methods & Instrumentation 8: 139–143. https://doi.org/10.3758/bf03201761.

Kapler T, Wright W (2004) Geo Time Information Visualization. In IEEE Symposium on Information Visualization, 25–32. Austin, TX: IEEE. https://doi.org/10.1109/infvis.2004.27.

Kelly M, Slingsby A, Dykes J et al (2013) Historical Internal Migration in Ireland. In Proceedings of GISRUK 2013. Liverpool, UK.

Kinkeldey C, MacEachren AM, Schiewe J (2014) How to Assess Visual Communication of Uncertainty? A Systematic Review of Geospatial Uncertainty Visualisation User Studies. The Cartographic Journal 51: 372–386. https://doi.org/10.1179/1743277414y.0000000099.

Kisilevich S, Krstajic M, Keim D et al (2010) Event-Based Analysis of People's Activities and Behavior Using Flickr and Panoramio Geotagged Photo Collections. In 14th International Conference on Information Visualisation, IV 2010, 289–296. London, UK: IEEE. https://doi.org/10.1109/iv.2010.94.

Konecny M (2011) Review: Cartography: Challenges and Potential in the Virtual Geographic Environments Era. Annals of GIS 17: 135–146. https://doi.org/10.1080/19475683.2011.602027.

Kounavis CD, Kasimati AE, Zamani ED (2012) Enhancing the Tourism Experience through Mobile Augmented Reality: Challenges and Prospects. International Journal of Engineering Business Management 4: article 10. https://doi.org/10.5772/51644.

Kourouthanassis PE, Boletsis C, Lekakos G (2015) Demystifying the Design of Mobile Augmented Reality Applications. Multimedia Tools and Applications 74: 1045–1066. https://doi.org/10.1007/s11042-013-1710-7.

Kraak MJ (2003a) Geovisualization Illustrated. ISPRS Journal of Photogrammetry and Remote Sensing 57: 390–399. https://doi.org/10.1016/s0924-2716(02)00167-3.

Kraak MJ (2003b) The Space–Time Cube Revisited from a Geovisualization Perspective. In Proceedings of the 21st International Cartographic Conference, 1988–1996. Durban, South Africa: ICA.

Kraak MJ (2008) Geovisualization and Time – New Opportunities for the Space–Time Cube. In Geographic Visualization: Concepts, Tools and Applications, edited by Martin Dodge, Martin Turner, and McDerby, 293–306. Chichester, West Sussex, UK: John Wiley & Sons. https://doi.org/10.1002/9780470987643.ch15.

Kurkovsky S, Koshy R, Novak V et al (2012) Current Issues in Handheld Augmented Reality. In 2012 International Conference on Communications and Information Technology (ICCIT), 68–72. Hammamet, Tunisia: IEEE. https://doi.org/10.1109/iccitechnol.2012.6285844.

Kveladze I, Kraak MJ, van Elzakker CPJM (2015) The Space-Time Cube as Part of a GeoVisual Analytics Environment to Support the Understanding of Movement Data. International Journal of Geographical Information Science 29: 2001–2016. https://doi.org/10.1080/13658816.2015.1058386.

Laney D (2001) 3D Data Management: Controlling Data Volume, Velocity, and Variety. META Group. https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf.

Li SN, Dragicevic S, Castro FA et al (2016) Geospatial Big Data Handling Theory and Methods: A Review and Research Challenges. ISPRS Journal of Photogrammetry and Remote Sensing 115: 119–133. https://doi.org/10.1016/j.isprsjprs.2015.10.012.

Li Y, Gong JH, Song YQ et al (2015) Design and Key Techniques of a Collaborative Virtual Flood Experiment That Integrates Cellular Automata and Dynamic Observations. Environmental Earth Sciences 74: 7059–7067. https://doi.org/10.1007/s12665-015-4716-9.

Li ZL, Huang PZ (2002) Quantitative Measures for Spatial Information of Maps. International Journal of Geographical Information Science 16: 699–709. https://doi.org/10.1080/13658810210149416.

Liang JM, Gong JH, Li Y (2015) Realistic Rendering for Physically Based Shallow Water Simulation in Virtual Geographic Environments (VGEs). Annals of GIS 21: 301–312. https://doi.org/10.1080/19475683.2015.1050064.

Liben LS, Downs RM (1993) Understanding Person-Space-Map Relations: Cartographic and Developmental Perspectives. Developmental Psychology 29: 739–752. https://doi.org/10.1037/0012-1649.29.4.739.

Lin H, Batty M, Jorgensen SE et al (2015) Virtual Environments Begin to Embrace Process-Based Geographic Analysis. Transactions in GIS 19: 493–498. https://doi.org/10.1111/tgis.12167.

Lin H, Chen M, Lu GN (2013a) Virtual Geographic Environment: A Workspace for Computer-Aided Geographic Experiments. Annals of the Association of American Geographers 103: 465–482. https://doi.org/10.1080/00045608.2012.689234.

Lin H, Chen M, Lu GN et al (2013b) Virtual Geographic Environments (VGEs): A New Generation of Geographic Analysis Tool. Earth-Science Reviews 126: 74–84. https://doi.org/10.1016/j.earscirev.2013.08.001.

Lin H, Gong JH (2001) Exploring Virtual Geographic Environments. Annals of GIS 7: 1–7. https://doi.org/10.1080/10824000109480550.

Lloyd D, Dykes J (2011) Human-Centered Approaches in Geovisualization Design: Investigating Multiple Methods Through a Long-Term Case Study. IEEE Transactions on Visualization and Computer Graphics 17: 2498–2507. https://doi.org/10.1109/tvcg.2011.209.

Lokka IE, Çöltekin A (2019) Toward Optimizing the Design of Virtual Environments for Route Learning: Empirically Assessing the Effects of Changing Levels of Realism on Memory. International Journal of Digital Earth 12: 137–155. https://doi.org/10.1080/17538947.2017.1349842.

Lokka IE, Çöltekin A, Wiener J et al (2018) Virtual Environments as Memory Training Devices in Navigational Tasks for Older Adults OPEN. Scientific Reports 8: 10809. https://doi.org/10.1038/s41598-018-29029-x.

Lü GN, Batty M, Strobl J et al (2019) Reflections and Speculations on the Progress in Geographic Information Systems (GIS): A Geographic Perspective. International Journal of Geographical Information Science 33 (2): 346–367. https://doi.org/10.1080/13658816.2018.1533136.

MacEachren AM (1982) Map Complexity: Comparison and Measurement. The American Cartographer 9: 31–46. https://doi.org/10.1559/152304082783948286.

MacEachren AM (1994) Visualization in Modern Cartography: Setting the Agenda. In Visualization in Modern Cartography, edited by Alan M. MacEachren and D. R. Fraser Taylor, 1–12. Oxford, UK: Pergamon Press.

MacEachren AM (1995) How Maps Work: Representation, Visualization, and Design. New York: Guilford Press.

MacEachren AM (2015) Visual Analytics and Uncertainty: It's Not About the Data. In EuroVis Workshop on Visual Analytics (EuroVA), 55–60. Cagliari, Sardinia, Italy. https://doi.org/10.2312/eurova.20151104.

MacEachren AM, Edsall R, Haug D et al (1999a) Exploring the Potential of Virtual Environments for Geographic Visualization. presented at the Annual Meeting of the Association of American Geographers, Honolulu, HI.

MacEachren AM, Jaiswal A, Robinson AC et al (2011) SensePlace2: Geo Twitter Analytics Support for Situational Awareness. In 2011 IEEE Conference on Visual Analytics, Science and Technology (VAST), 181–190. Providence, RI: IEEE. https://doi.org/10.1109/vast.2011.6102456.

MacEachren AM, Kraak MJ (2001) Research Challenges in Geovisualization. Cartography and Geographic Information Science 28: 3–12. https://doi.org/10.1559/152304001782173970.

MacEachren AM, Kraak MJ, Verbree E (1999b) Cartographic Issues in the Design and Application of Geospatial Virtual Environments. In Proceedings of the 19th International Cartographic Conference. Ottawa, Canada: ICA.

MacEachren AM, Roth RE, O'Brien J et al (2012) Visual Semiotics & Uncertainty Visualization: An Empirical Study. IEEE Transactions on Visualization and Computer Graphics 18: 2496–2505. https://doi.org/10.1109/tvcg.2012.279.

Marriott K, Schreiber F, Dwyer T et al (2018) Immersive Analytics. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-030-01388-2.

Mekni M (2010) Hierarchical Path Planning for Situated Agents in Informed Virtual Geographic Environments. In SIMUTools 2010. Torremolinos, Malaga, Spain. https://doi.org/10.4108/icst.simutools2010.8811.
Mellado N, Vanderhaeghe D, Hoarau C et al (2017) Constrained Palette-Space Exploration. ACM Transactions on Graphics 36: 1–14. https://doi.org/10.1145/3072959.3073650.
Milgram P, Kishino F (1994) A Taxonomy of Mixed Reality Visual Displays. IEICE TRANSACTIONS on Information and Systems 77: 1321–1329.
Monmonier M (2018) How to Lie with Maps (Third Edition). Chicago, IL: University of Chicago Press.
Newcombe N, Bandura MM, Taylor DG (1983) Sex Differences in Spatial Ability and Spatial Activities. Sex Roles 9: 377–386. https://doi.org/10.1007/bf00289672.
Nöllenburg M (2007) Human-Centered Visualization Environments, GI-Dagstuhl Research Seminar, Dagstuhl Castle, Germany, March 5–8, 2006, Revised Lectures 4417: 257–294. https://doi.org/10.1007/978-3-540-71949-6_6.
Olsson T (2012) User Expectations, Experiences of Mobile Augmented Reality Services. PhD Thesis, Tampere, Finland: Tampere University of Technology. https://tutcris.tut.fi/portal/files/5450806/olsson.pdf.
Ooms K, Maeyer PD, Fack V (2015) Listen to the Map User: Cognition, Memory, and Expertise. The Cartographic Journal 52: 3–19. https://doi.org/10.1179/1743277413y.0000000068.
Oprean D, Wallgrün JO, Duarte JMP et al (2018) Developing and Evaluating VR Field Trips. In Proceedings of Workshops and Posters of the 13th International Conference on Spatial Information Theory (COSIT 2017), edited by Eliseo Clementini, Eliseo Fogliaroni, and Andrea Ballatore, 105–110. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-63946-8_22.
Oprean D, Simpson M, Klippel A (2017) Collaborating remotely: An evaluation of immersive capabilities on spatial experiences and team membership. International Journal of Digital Earth, 11(4): 420–436. https://doi.org/10.1080/17538947.2017.1381191.
Patterson T, Kelso NV (2004) Hal Shelton Revisited: Designing and Producing Natural-Color Maps with Satellite Land Cover Data. Cartographic Perspectives 47: 28–55. https://doi.org/10.14714/cp47.470.
Perkins C (2008) Cultures of Map Use. The Cartographic Journal 45: 150–158. https://doi.org/10.1179/174327708x305076.
Petrasova A, Harmon B, Petras V et al (2015) Tangible Modeling with Open Source GIS. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-25775-4.
Pettit C, Lieske SN, Jamal M (2017a) CityDash: Visualising a Changing City Using Open Data. In Planning Support Science for Smarter Urban Futures, edited by Stan Geertman, Allan Andrews, Chris Pettit, and John Stillwell, 337–353. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-57819-4_19.
Pettit CJ, Leao SZ (2017) Dashboard. In Encyclopedia of Big Data, edited by Laurie A. Schintler and Connie L. McNeely, 1–6. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-32001-4.
Pettit CJ, Russel ABM, Michael A et al (2010) Realising an EScience Platform to Support Climate Change Adaptation in Victoria. In 2010 IEEE Sixth International Conference on E-Science, 73–80. IEEE. https://doi.org/10.1109/escience.2010.32.
Pettit C, Tice A, Randolph B (2017b) Using an Online Spatial Analytics Workbench for Understanding Housing Affordability in Sydney. In Seeing Cities Through Big Data, Research, Methods and Applications in Urban Informatics, edited by Piyushimita Thakuriah, Nebiyou Tilahun, and Moira Zellner, 233–255. Cham: Springer. https://doi.org/10.1007/978-3-319-40902-3_14.
Pindat C, Pietriga E, Chapuis O et al (2012) JellyLens: Content-Aware Adaptive Lenses. In UIST '12 Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology, 261–270. Cambridge, MA: ACM. https://doi.org/10.1145/2380116.2380150.

Pirolli P, Card S (2005) The Sensemaking Process and Leverage Points for Analyst Technology as Identified through Cognitive Task Analysis. In Proceedings of International Conference on Intelligence Analysis. Maclean, VA.
Priestnall G, Jarvis C, Burton A et al (2012) Virtual Geographic Environments. In Teaching Geographic Information Science and Technology in Higher Education, edited by David J. Unwin, Kenneth E. Foote, Nicholas J. Tate, and David DiBiase, 257–288. Chichester, West Sussex, UK: John Wiley & Sons. https://doi.org/10.1002/9781119950592.ch18.
Rautenbach V, Çöltekin A, Coetzee S (2015) Exploring the Impact of Visual Complexity Levels in 3D City Models on the Accuracy of Individuals' Orientation and Cognitive Maps. In ISPRS Geospatial Week 2015, Workshop II-3/W5, edited by Sidonie Christophe and Arzu Çöltekin, 499–506. La Grande Motte, France. https://doi.org/10.5194/isprsannals-ii-3-w5-499-2015.
Richter KF, Tomko M, Çöltekin A (2015) Are We There Yet? Spatial Cognitive Engineering for Situated Human-Computer Interaction. In CESIP 2015: Cognitive Engineering for Spatial Information Processes: From User Interfaces to Model-Driven Design. Workshop at COSIT 2015, 1–7. Santa Fe, NM.
Rink K, Chen C, Bilke L et al (2018) Virtual Geographic Environments for Water Pollution Control. International Journal of Digital Earth 11: 397–407. https://doi.org/10.1080/17538947.2016.1265016.
Roberts JC (2007) State of the Art: Coordinated & Multiple Views in Exploratory Visualization. In Fifth International Conference on Coordinated and Multiple Views in Exploratory Visualization (CMV 2007), 61–71. Zürich, Switzerland: IEEE. https://doi.org/10.1109/cmv.2007.20.
Robertson G, Fernandez R, Fisher D et al (2008) Effectiveness of Animation in Trend Visualization. IEEE Transactions on Visualization and Computer Graphics 14: 1325–1332. https://doi.org/10.1109/tvcg.2008.125.
Robertson GG, Mackinlay JD (1993) The Document Lens. In UIST '93 Proceedings of the 6th Annual ACM Symposium on User Interface Software and Technology, 101–108. Atlanta, GA: ACM. https://doi.org/10.1145/168642.168652.
Robinson A (2017) Geovisual Analytics. In The Geographic Information Science & Technology Body of Knowledge (3rd Quarter 2017 Edition), edited by John P. Wilson. UCGIS. https://doi.org/10.22224/gistbok/2017.3.6.
Robinson AC, Chen J, Lengerich EJ et al (2005) Combining Usability Techniques to Design Geovisualization Tools for Epidemiology. Cartography and Geographic Information Science 32: 243–255. https://doi.org/10.1559/152304005775194700.
Robinson AC, Demšar U, Moore AB et al (2017) Geospatial Big Data and Cartography: Research Challenges and Opportunities for Making Maps That Matter. International Journal of Cartography 3: 32–60. https://doi.org/10.1080/23729333.2016.1278151.
Roth RE, Çöltekin A, Delazari L et al (2017) User Studies in Cartography: Opportunities for Empirical Research on Interactive Maps and Visualizations. The International Journal of Cartography 3: 61–89. https://doi.org/10.1080/23729333.2017.1288534.
Ruas A, Perret J, Curie F et al (2011) Conception of a GIS-Platform to Simulate Urban Densification Based on the Analysis of Topographic Data. In Advances in Cartography and GIScience, Volume 2, edited by Anne Ruas, 413–430. Berlin Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-642-19214-2.
Russo P, Pettit C, Çöltekin A et al (2013) Understanding Soil Acidification Process Using Animation and Text: An Empirical User Evaluation With Eye Tracking. In Proceedings of the 26th International Cartographic Conference, 431–448. Berlin Heidelberg: Springer-Verlag. https://doi.org/10.1007/978-3-642-32618-9_31.
Sachs J, Schmidt-Traub G, Kroll C et al (2018). New York: Bertelsmann Stiftung and Sustainable Development Solutions Network (SDSN).
Scaife M, Rogers Y (1996) External Cognition: How Do Graphical Representations Work? International Journal of Human-Computer Studies 45: 185–213. https://doi.org/10.1006/ijhc.1996.0048.

Schnur S, Bektaş K, Çöltekin A (2018) Measured and Perceived Visual Complexity: A Comparative Study among Three Online Map Providers. Cartography and Geographic Information Science 45: 238–254. https://doi.org/10.1080/15230406.2017.1323676.

Schnur S, Bektaş K, Salahi M et al (2010) A Comparison of Measured and Perceived Visual Complexity for Dynamic Web Maps. In Proceedings of GIScience 2010, edited by Ross Purves and Robert Weibel, 1–4. Zürich, Switzerland. http://www.giscience2010.org/pdfs/paper_181.pdf.

Semmo A, Döllner J (2014) An Interaction Framework for Level-of-Abstraction Visualization of 3D Geovirtual Environments. In MapInteract'14, 43–49. Dallas/Fort Worth, Texas: ACM. https://doi.org/10.1145/2677068.2677072.

Semmo A, Trapp M, Kyprianidis JE et al (2012) Interactive Visualization of Generalized Virtual 3D City Models Using Level-of-Abstraction Transitions. Computer Graphics Forum 31: 885–894. https://doi.org/10.1111/j.1467-8659.2012.03081.x.

Shen S, Gong JH, Liang JM et al (2018) A Heterogeneous Distributed Virtual Geographic Environment—Potential Application in Spatiotemporal Behavior Experiments. ISPRS International Journal of Geo-Information 7: 54. https://doi.org/10.3390/ijgi7020054.

Sherman WR, Craig AB (2003) Understanding VR: Understanding Virtual Reality: Interface, Application, and Design. San Francisco: Morgan Kaufmann Publishers, Inc.

Sinnott RO, Bayliss C, Bromage A et al (2015) The Australia Urban Research Gateway. Concurrency and Computation: Practice and Experience 27: 358–375. https://doi.org/10.1002/cpe.3282.

Skupin A, Buttenfield BP (1997) Spatial Metaphors for Visualizing Information Spaces. Accounting, Organizations and Society 32: 649–667. https://doi.org/10.1016/j.aos.2007.02.001.

Slater M (2009) Place Illusion and Plausibility Can Lead to Realistic Behaviour in Immersive Virtual Environments. Philosophical Transactions of the Royal Society B: Biological Sciences 364: 3549–3557. https://doi.org/10.1098/rstb.2009.0138.

Slingsby A (2018) Tilemaps for Summarising Multivariate Geographical Variation. In Paper Presented at VIS 2018. Berlin, Germany: IEEE.

Slingsby A, van Loon E (2016) Exploratory Visual Analysis for Animal Movement Ecology. Computer Graphics Forum 35 (3): 471–480. https://doi.org/10.1111/cgf.12923.

Slingsby A, Wood J, Dykes J (2010) Treemap Cartography for Showing Spatial and Temporal Traffic Patterns. Journal of Maps 6: 135–146. https://doi.org/10.4113/jom.2010.1071.

Slocum TA, Blok C, Jiang B et al (2001) Cognitive and Usability Issues in Geovisualization. Cartography and Geographic Information Science 28: 61–75. https://doi.org/10.1559/152304001782173998.

Slocum TA, Cliburn DC, Feddema JJ et al (2003) Evaluating the Usability of a Tool for Visualizing the Uncertainty of the Future Global Water Balance. Cartography and Geographic Information Science 30 (4): 299–317. https://doi.org/10.1559/152304003322606210.

Slocum TA, McMaster RB, Kessler FC et al (2008) Thematic Cartography and Geovisualization (3rd Edition). Upper Saddle Hall, NJ: Prentice Hall.

Smallman HS, John MS (2005) Naive Realism: Misplaced Faith in Realistic Displays. Ergonomics in Design: The Quarterly of Human Factors Applications 13: 6–13. https://doi.org/10.1177/106480460501300303.

Spekat A, Kreienkamp F (2007) Somewhere over the Rainbow – Advantages and Pitfalls of Colourful Visualizations in Geosciences. Advances in Science and Research 1: 15—21. https://doi.org/10.5194/asr-1-15-2007.

Spence R (2007) Information Visualization: Design for Interaction (2nd Edition). Harlow, Essex, UK: Pearson Education Limited.

Thomas JJ, Cook KA (2005) Illuminating the Path: The Research and Development Agenda for Visual Analytics. New York: IEEE.

Thoresen JC, Francelet R, Coltekin A et al (2016) Not All Anxious Individuals Get Lost: Trait Anxiety and Mental Rotation Ability Interact to Explain Performance in Map-Based Route Learning in Men. Neurobiology of Learning and Memory 132: 1–8. https://doi.org/10.1016/j.nlm.2016.04.008.

Thyng K, Greene C, Hetl R et al (2016) True Colors of Oceanography: Guidelines for Effective and Accurate Colormap Selection. Oceanography 29: 9–13. https://doi.org/10.5670/oceanog.2016.66.

Tobler WR (1987) Experiments In Migration Mapping By Computer. Cartography and Geographic Information Science 14 (2): 155–163. https://doi.org/10.1559/152304087783875273.

Tomaszewski B, MacEachren AM (2012) Geovisual Analytics to Support Crisis Management: Information Foraging for Geo-Historical Context. Information Visualization 11: 339–359. https://doi.org/10.1177/1473871612456122.

Tominski C, Gladisch S, Kister U et al (2017) Interactive Lenses for Visualization: An Extended Survey. Computer Graphics Forum 36: 173–200. https://doi.org/10.1111/cgf.12871.

Tominski C, Schumann H, Andrienko G et al (2012) Stacking-Based Visualization of Trajectory Attribute Data. IEEE Transactions on Visualization and Computer Graphics 18: 2565–2574. https://doi.org/10.1109/tvcg.2012.265.

Tomko M, Winter S (2019) Beyond Digital Twins-A Commentary. Environment and Planning B in press. https://doi.org/10.1177/2399808318816992.

Torrens PM (2015) Slipstreaming Human Geosimulation in Virtual Geographic Environments. Annals of GIS 21: 325–344. https://doi.org/10.1080/19475683.2015.1009489.

Touya G, Hoarau C, Christophe S (2016) Clutter and Map Legibility in Automated Cartography: A Research Agenda. Cartographica: The International Journal for Geographic Information and Geovisualization 51: 198–207. https://doi.org/10.3138/cart.51.4.3132.

Tsai TH, Chang HT, Yu MC et al (2016) Design of a Mobile Augmented Reality Application: An Example of Demonstrated Usability. In UAHCI 2016, 198–205. Cham, Switzerland: Springer International Publishing. https://doi.org/10.1007/978-3-319-40244-4_19.

Tufte ER (1983) The Visual Display of Quantitative Information. Cheshire, CT: Graphics Press.

Tversky B, Morrison JB and Betrancourt M (2002) Animation: Can It Facilitate? International Journal of Human-Computer Studies 57: 247–262. https://doi.org/10.1006/ijhc.2002.1017.

United Nations Population Division (2018) World Urbanization Prospects: The 2018 Revision. New York: United Nations.

Uttal DH, Meadow NG, Tipton E et al (2013) The Malleability of Spatial Skills: A Meta-Analysis of Training Studies. Psychological Bulletin 139: 352–402. https://doi.org/10.1037/a0028446.

van Wijk JJ, Nuij WAA (2003) Smooth and Efficient Zooming and Panning. In INFOVIS'03 Proceedings of the Ninth Annual IEEE Conference on Information Visualization, 15–22. Seattle, WA.

Vert S, Dragulescu B, Vasiu R (2014) LOD4AR: Exploring Linked Open Data with a Mobile Augmented Reality Web Application. In 13th International Semantic Web Conference (ISWC 2014), 185–188. Trentino, Italy.

Viard T, Caumon G, Lévy B (2011) Adjacent versus Coincident Representations of Geospatial Uncertainty: Which Promote Better Decisions? Computers & Geosciences 37 (4): 511–520. https://doi.org/10.1016/j.cageo.2010.08.004.

Voinov A, Çöltekin A, Chen M et al (2018) Virtual Geographic Environments in Socio-Environmental Modeling: A Fancy Distraction or a Key to Communication? International Journal of Digital Earth 11: 408–419. https://doi.org/10.1080/17538947.2017.1365961.

Weichelt B, Yoder A, Bendixsen C et al (2018) Augmented Reality Farm MAPPER Development: Lessons Learned from an App Designed to Improve Rural Emergency Response. Journal of Agromedicine 23: 284–296. https://doi.org/10.1080/1059924x.2018.1470051.

Wen YN, Chen M, Lu GN et al (2013) Prototyping an Open Environment for Sharing Geographical Analysis Models on Cloud Computing Platform. International Journal of Digital Earth 6: 356–382. https://doi.org/10.1080/17538947.2012.716861.

Wen YN, Chen M, Yue SS et al (2017) A Model-Service Deployment Strategy for Collaboratively Sharing Geo-Analysis Models in an Open Web Environment. International Journal of Digital Earth 10: 405–425. https://doi.org/10.1080/17538947.2015.1131340.

Wickham H, Hofmann H, Wickham C et al (2012) Glyph-Maps for Visually Exploring Temporal Patterns in Climate Data and Models. Environmetrics 23: 382–393. https://doi.org/10.1002/env.2152.

Widjaja I, Russo P, Pettit C et al (2014) Modeling Coordinated Multiple Views of Heterogeneous Data Cubes for Urban Visual Analytics. International Journal of Digital Earth 8: 558–578. https://doi.org/10.1080/17538947.2014.942713.

Wood J, Slingsby A, Dykes J (2011) Visualizing the Dynamics of London's Bicycle-Hire Scheme. Cartographica: The International Journal for Geographic Information and Geovisualization 46: 239–251. https://doi.org/10.3138/carto.46.4.239.

Wu HK, Lee WY, Chang HY et al (2013) Current Status, Opportunities and Challenges of Augmented Reality in Education. Computers & Education 62: 41–49. https://doi.org/10.1016/j.compedu.2012.10.024.

Wüest R, Nebiker S (2018) Geospatial Augmented Reality for the Interactive Exploitation of Large-Scale Walkable Orthoimage Maps in Museums. Proceedings of the ICA 1: 1–6. https://doi.org/10.5194/ica-proc-1-124-2018.

Xu BL, Lin H, Chiu LS et al (2011) Collaborative Virtual Geographic Environments: A Case Study of Air Pollution Simulation. Information Sciences 181: 2231–2246. https://doi.org/10.1016/j.ins.2011.01.017.

Yang YL, Dwyer T, Jenny B et al (2019) Origin-Destination Flow Maps in Immersive Environments. IEEE Transactions on Visualization and Computer Graphics 25: 693–703. https://doi.org/10.1109/tvcg.2018.2865192.

Yue SS, Chen M, Wen YN et al (2016) Service-Oriented Model-Encapsulation Strategy for Sharing and Integrating Heterogeneous Geo-Analysis Models in an Open Web Environment. ISPRS Journal of Photogrammetry and Remote Sensing 114: 258–273. https://doi.org/10.1016/j.isprsjprs.2015.11.002.

Zhang F, Hu MY, Che WT et al (2018) Framework for Virtual Cognitive Experiment in Virtual Geographic Environments. ISPRS International Journal of Geo-Information 7: 36. https://doi.org/10.3390/ijgi7010036.

Zheng PB, Tao H, Yue SS et al (2017) A Representation Method for Complex Road Networks in Virtual Geographic Environments. ISPRS International Journal of Geo-Information 6: 372. https://doi.org/10.3390/ijgi6110372.

Zhu J, Zhang H, Yang XF et al (2015) A Collaborative Virtual Geographic Environment for Emergency Dam-Break Simulation and Risk Analysis. Journal of Spatial Science 61: 133–155. https://doi.org/10.1080/14498596.2015.1051148.

Arzu Çöltekin is an Associate Professor at the Institute of Interactive Technologies of University of Applied Sciences and Arts Northwestern Switzerland. Her background is in geographic information science and engineering; and her research interests are at the intersection of visuospatial cognition, human-computer interaction, mixed reality and GIScience. She is a council member with the ISDE, chairs ISPRS Working Group on Geovisualization, Augmented and Virtual Reality; and is a co-chair of ICA Commission on Visual Analytics.

Amy L. Griffin is a Senior Lecturer in the School of Science at RMIT University in Melbourne, Australia. Her research is focused on understanding how perceptual, cognitive and emotional processes shape how people read maps and think with spatial information. She is currently a co-chair of the ICA Commission on Cognitive Issues in Geographic Information Visualization.

Aidan Slingsby is a Senior Lecturer at giCentre, Department of Computer Science, City, University of London. He has a background in GIS, but now works more generally in information visualisation and visual analytics. He is interested in the role of data visualization in the analysis of data, particularly those that are spatial and temporal.

Anthony C. Robinson is an Associate Professor of Geography at Penn State University. His science focuses on improving the utility of geospatial information through cartographic design and user evaluation. Dr. Robinson is Assistant Director of the GeoVISTA Center and Director of Online Geospatial Education at Penn State, and co-chairs the Commission on Visual Analytics for the International Cartographic Association.

Sidonie Christophe is a Senior Researcher in Geographic Information Sciences at the Paris-Est University, at the research Laboratory in Sciences and Technologies in Geographic Information. She focuses on the design, the perception and the uses of graphic representations for geovisualization, augmented and mixed reality and visual spatio-temporal analysis of urban and environmental dynamics. She is the co-chair of the ISPRS working group on "Geovisualization, VR & AR", and the co-chair of the ICA Commission on "Cognitive Issues in Geovisualization".

Victoria Rautenbach is a Senior Lecturer in the Department of Geography, Geoinformatics and Meteorology at the University of Pretoria, South Africa. Victoria is the scientific research secretary of the ISPRS WG IV/9 Geovisualization, Augmented and Virtual Reality. She is passionate about open source, open education and open data. Victoria currently serves on the OSGeo Board of Directors and the co-chair of GeoForAll.

Min Chen is a Professor and the Vice Director of Key Lab of Virtual Geographic Environment (Ministry of Education of PRC), Nanjing Normal University, China. His research interests are geographic modeling and simulation, virtual geographic environment. He is the executive editor of Annals of GIS and an editor of Environmental Modelling & Software.

Christopher Pettit is a Professor and the Director of City Analytics, UNSW Sydney. His expertise is in the convergence of the fields of city planning and digital technologies including Geographical Information Sciences (GIS). He is currently the Chair of the Board for CUPUM (Computer in Urban Planning and Urban Management).

Alexander Klippel is a Professor in Geography and Information Sciences at Penn State and the inaugural director of the Center of Immersive Experiences. This research interests lie at the interface of technology and cognition and address questions such as immersive learning, immersive decision-making, and the centrality of space for cognitive processes.

Chapter 8
Transformation in Scale for Continuous Zooming

Zhilin Li and Haowen Yan

Abstract This chapter summarizes the theories and methods in continuous zooming for Digital Earth. It introduces the basic concepts of and issues in continuous zooming and transformation in scale (or multiscale transformation). It presents the theories of transformation in scale, including the concepts of multiscale versus variable scale, transformation in the Euclidean space versus the geographical space, and the theoretical foundation for transformation in scale, the Natural Principle. It addresses models for transformations in scale, including space-primary hierarchical models, feature-primary hierarchical models, models of transformation in scale for irregular triangulation networks, and the models for geometric transformation of map data. It also discusses the mathematical solutions to transformations in scale (including upscaling and downscaling) for both raster (numerical and categorical data) and vector (point set data, line data set and area data) data. In addition, some concluding remarks are provided.

Keywords Continuous zooming · Transformation in scale · Natural principle · Multiscale · Variable scale

Z. Li (✉)
Department of Land Surveying and Geo-Informatics, Hong Kong Polytechnic University, Hong Kong, China
e-mail: zl.li@polyu.edu.hk

H. Yan
Department of Geographic Information Science, Lanzhou Jiaotong University, Lanzhou, China

Z. Li
State-Province Joint Engineering Laboratory of Spatial Information Technology for High-Speed Railway Safety, Southwest Jiaotong University, Chengdu, China

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_8

8.1 Continuous Zooming and Transformation in Scale: An Introduction

8.1.1 Continuous Zooming: Foundation of the Digital Earth

Continuous zooming is a fundamental function of a Digital Earth, as the demand for such a function has been vividly portrayed by then-US Vice President Al Gore in his famous speech “The Digital Earth: Understanding Our Planet in the twenty-first Century” (Gore 1998):

> Imagine, for example, a young child going to a Digital Earth exhibit at a local museum. After donning a head-mounted display, she sees Earth as it appears from space. Using a data glove, she zooms in, using higher and higher levels of resolution, to see continents, then regions, countries, cities, and finally individual houses, trees, and other natural and man-made objects.

The cascade scene seen by the young child is a result of continuous zooming. Such zooming can be realized by continuously displaying a series of Earth images taken at a given position and changing the focal length of the camera lens continuously or displaying images taken at different heights continuously but with at a fixed camera focal length.

In theory, to make the display visually smooth, the differences between two images should be sufficiently small, thus the number of images in such a series is very large, which demands huge data storage. Thus, it is a very difficult, if not impossible, problem.

8.1.2 Transformation in Scale: Foundation of Continuous Zooming

In practice, Earth images are acquired and stored at discrete scales (e.g., 1:500,000, 1:100,000, 1:10,000) or different resolutions (e.g., 100, 10, 1, 0.5 m), leading to the term *multiscale representation*. Figure 8.1 shows a series of satellite images covering Hong Kong Polytechnic University at six different scales, extracted from Google Maps. If such images at discrete scales are displayed in sequence, there will be a visual jump between two images. The obviousness of the visual jump is dependent on the magnitude of the scale difference. The smaller the difference between the two scales is, the less apparent the visual jump will be.

To minimize the effect of such visual jumps, some techniques are required to smooth the transformations from one scale to another scale to make the display appear like continuous zooming. This transformation in scale is the foundation of continuous zooming. Thus, transformation in scale, also called multiscale transformation, is the topic of this chapter.

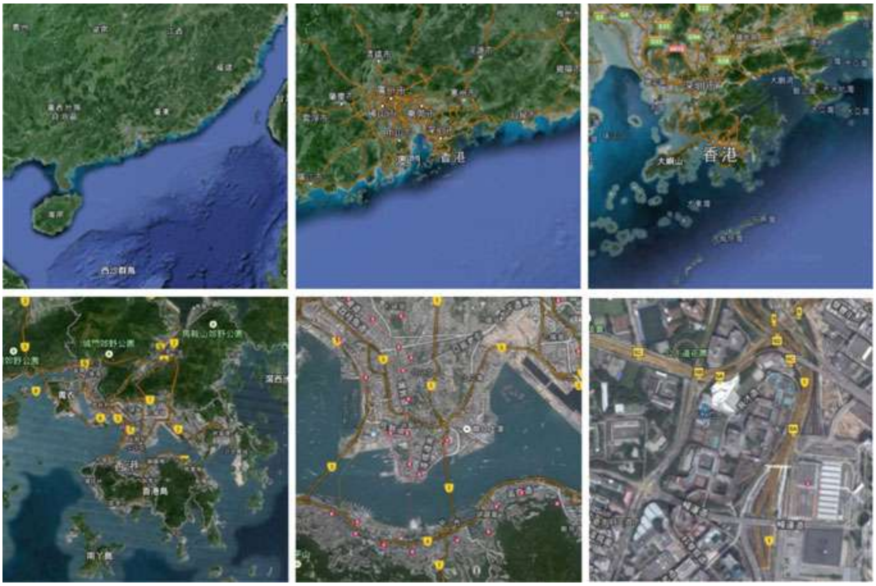


Fig. 8.1 A series of images covering HK Polytechnic University at different scales (from Google Maps)

8.1.3 Transformation in Scale: A Fundamental Issue in Disciplines Related to Digital Earth

Transformation in scale is one of the most important but unsolved issues in various disciplines related to Digital Earth, such as mapping, geography, geomorphology, oceanography, soil science, social sciences, hydrology, environmental sciences and urban studies. Typical examples are map generalization and the modifiable areal unit problem (MAUP). Although transformation in scale is a traditional topic, it has been a critical issue in this digital era.

Transformation in scale has attracted attention from disciplines related to Digital Earth since the 1980s because a few important publications on the scale issue in that period awakened researchers in relevant areas. Openshaw (1984) revisited the *MAUP*. Abler (1987) reported that multiscale representation was identified as one of the initiatives of the National Center for Geographic Information and Analysis (NCGIA), and noted that zooming and overlay are the two most exciting functions in a geographical information system. Since then, the scale issue has been included in many research agendas (e.g., Rhind 1988; UCGIS 2006) and has become popular in the geo-information community.

The first paper on the scale issue in remote sensing was also published in 1987 (Woodcock and Strahler 1987). Later, in 1993, the issue of scaling from point to regional- or global-scale estimates of the surface energy fluxes attracted great attention at the Workshop on Thermal Remote Sensing held at La Londe les Maures, France from September 20–24. Scale became a hot topic in remote sensing as well.

As a result, many papers on the scale issue have been published in academic journals and at conferences related to Digital Earth. Other papers have been p in the form of edited books, such as *Scaling Up in Hydrology Using Remote Sensing* edited by Stewart et al. (1996), *Scale in Remote Sensing and GIS* edited by Quattrochi and Goodchild (1997), *Scale Dependence and Scale Invariance in Hydrology* edited by Sposito (1998), *Modelling Scale in Geographical Information Science* edited by Tate and Atkinson (2001), *Scale and Geographic Inquiry: Nature, Society and Method* edited by Sheppard and McMaster (2004), *Generalisation of Geographic Information: Cartographic Modelling and Applications* edited by Mackaness et al. (2007), and *Scale Issues in Remote sensing* edited by Weng (2014). Authored research monographs have also been published by researchers, e.g., *Algorithmic Foundation of Multi-Scale Spatial Representation* by Li (2007) and *Integrating Scale in Remote Sensing and GIS* by Zhang et al. (2017).

8.2 Theories of Transformation in Scale

Transformation in scale is the modeling of spatial data or spatial representations from one scale to another by employing mathematical models and/or algorithms developed based on certain scaling theories and/or principles. This section describes such scaling theories and/or principles.

8.2.1 Transformation in Scale: Multiscale Versus Variable Scale

To facilitate zooming, not necessarily continuous, a common practice of service providers such as Google Maps, Virtual Earth and Tianditu is to organize maps and images into nearly 20 levels (scales or resolutions), from global level to street level. Figure 8.2 shows a series of maps covering Hong Kong Polytechnic University at six different scales (extracted from Google Maps). This follows the tradition of organizing maps by national map agencies. For example, the United States Geological Survey (USGS) produces topographic maps at scales of 1:500,000, 1:250,000, 1:100,000, 1:50,000 and 1:24,000; the Chinese State Bureau of Surveying and Mapping produces maps at scales of 1:4,000,000, 1:1,000,000, 1:250,000, 1:50,000 and 1:10,000; the Ordnance Survey of the UK produces maps at scales of 1:50,000, 1:25,000 and 1:10,000; and the German federal states produces maps at 1:1,000,000, 1:250,000, 1:100,000, 1:50,000, 1:25,000 and 1:10,000 scales. These maps at different scales contain information at different levels of detail, and thus are suitable for different

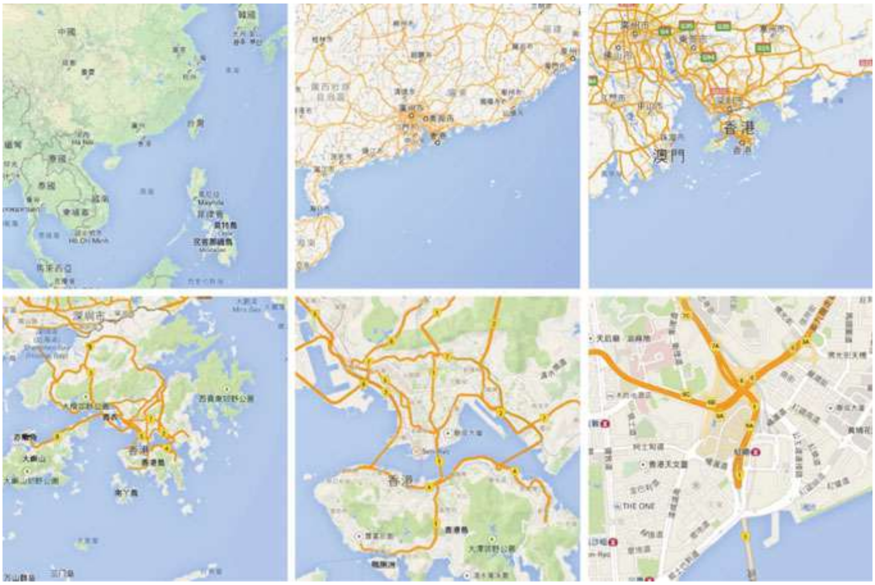


Fig. 8.2 A series of maps covering Hong Kong Polytechnic University at different scales (extracted from Google Maps)

applications. Such a scale is also called the *cartographic ratio*. Similarly, image data and digital elevation models (DEMs) are also produced and stored at discrete scales. In these two cases, the scale is normally indicated by *resolution*.

This kind of representation is called multiscale representation. In such cases, the cartographic ratio is uniform across a map and/or an image. Thus, such representations have multiple cartographic ratios. The cartographic ratio may vary across a representation (e.g., oblique view), leading to the term variable scale representation; the resolution may also vary across a representation, leading to the term variable resolution representation. As a result, the term multiscale might mean different things to different people, i.e., multi cartographic ratio, variable cartographic ratio, multi resolution and variable resolution. This leads to nine different kinds of transformations in scale, as shown in Fig. 8.3.

8.2.2 Transformations in Scale: Euclidean Versus Geographical Space

In *Euclidean space*, an increase in scale will commonly cause an increase in length, area and volume; and a decrease in scale will cause a decrease in length, area and volume, accordingly. Figure 8.4 shows an example of scale reduction and increase in

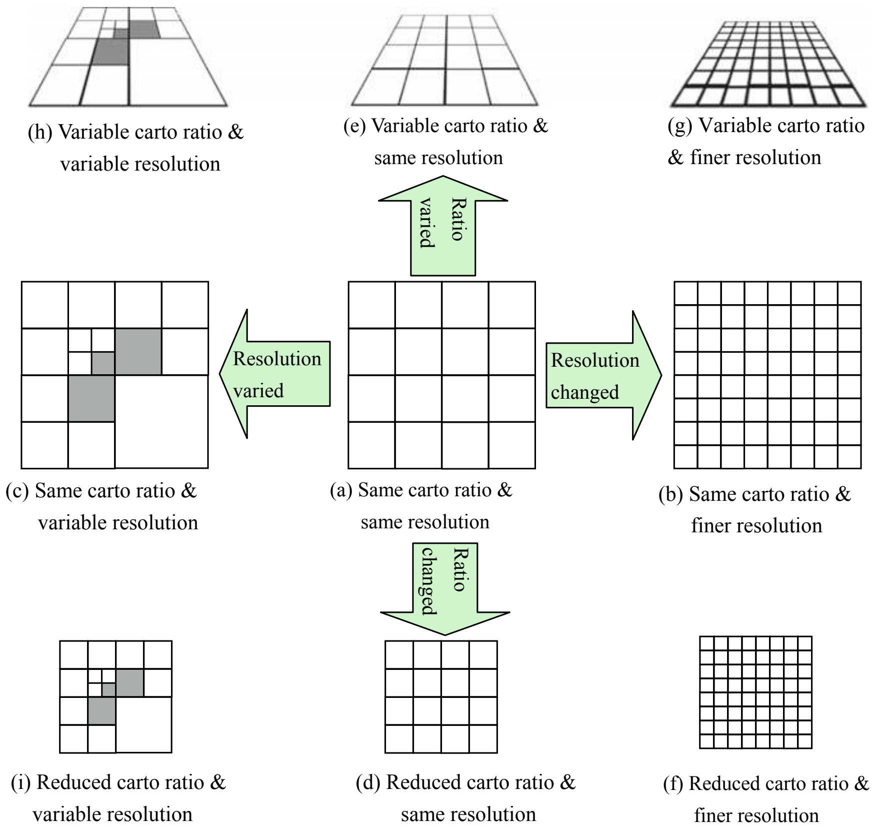


Fig. 8.3 Nine types of transformations in scale (Li 2008)

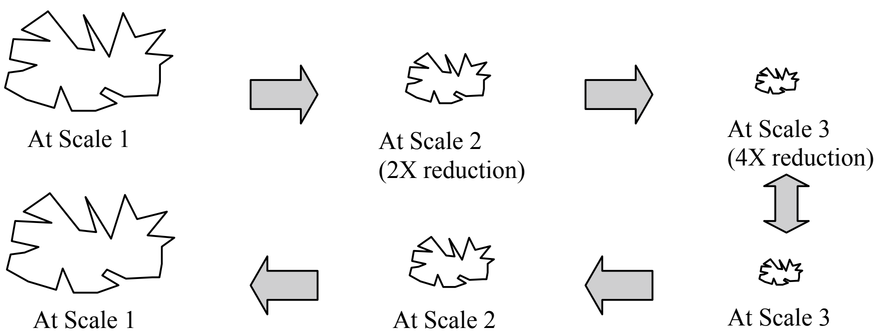


Fig. 8.4 Scale change in Euclidean space: a reversible process (Li 2007)

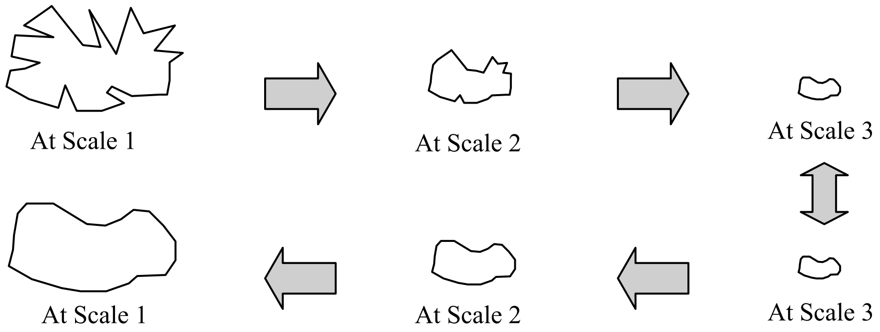


Fig. 8.5 Scale change in 2D geographical space: lost complexity is not recoverable (Li 2007)

a 2D Euclidean space. In such a transformation in scale, the absolute complexity of a feature or features remains unchanged. That is, the transformations are reversible.

However, the geographical space *is fractal*. If one measures a coastal line using different measurement units, then different lengths will be obtained. The smaller the measurement unit is, the longer the length obtained. Similarly, different length values will be obtained when measuring a coastal line represented on maps at different scales using identical measurement units at map scale. That is, the transformation in scale in fractal geographical space is quite different from that in Euclidean space.

For a given area on a terrain surface, the size of the graphic representation (or map space) on a smaller scale map is reduced compared with that on larger scale maps. The complexity of the graphics on a smaller scale map remain compatible with larger scale maps. However, the absolute complexity is reduced. As a result, if the graphics on a smaller map are enlarged back to the size on the larger scale map, the level of complexity of the enlarged representation will appear to be reduced. Figure 8.5 illustrates such a case. In a fractal geographical space, the level of complexity cannot be recovered by an increase in scale. In other words, the transformations in scale in such a geographical space are not reversible.

The transformation in scale is also termed *scaling*. The process of making the resolution coarser (or making the map scale smaller) is called *upscaling*. In contrast, the transformation process to make the resolution finer (or map scale larger) is called *downscaling*.

8.2.3 Theoretical Foundation for Transformation in Scale: The Natural Principle

One question that arises is "does such a transformation follow any principle or law?" The answer is "yes". Li and Openshaw (1993) formulated the *Natural Principle* for such a transformation in scale in fractal geographical space.

Li and Openshaw (1993) made use of the terrain surface viewed from different height levels as an example to illustrate the *Natural Principle*, as follows:

- When one views the terrain surface from the Moon, all terrain variations disappear, and one can only see a blue ball;
- When one views the terrain surface from a satellite, then the terrain surface becomes visible, but the terrain surface looks very smooth;
- When one views the terrain surface from an airplane, the main characteristics of the terrain variations become very clear, but small details do not appear; and
- When one views the terrain surface from a position on ground, the main characteristics of the terrain variations become lost, and one sees small details.

When the viewpoint is higher, the ground area corresponding to the human eyes' resolution becomes larger, but all detailed variations within this ground area can no longer be seen, and thus the terrain surface appears more abstract. These examples underline a universal principle, the *Natural Principle* as termed by Li and Openshaw (1993). It can be stated as follows:

> For a given scale of interest, all details about the spatial variations of geographical objects (features) beyond a certain limitation cannot be presented and can thus be neglected.

It follows that a simple corollary to this process can be used as a basis for transformations in scale. The corollary can be stated as follows (Li and Openshaw 1993):

> By using a criterion similar to the limitation of human eyes' resolution, and, neglecting all the information about the spatial variation of spatial objects (features) beyond this limitation, zooming (or generalization) effects can be achieved.

Li and Openshaw (1992) also term such a limitation as the smallest visible object (SVO) or smallest visible size (SVS) in other literature (Li 2007). Figure 8.6 illustrates

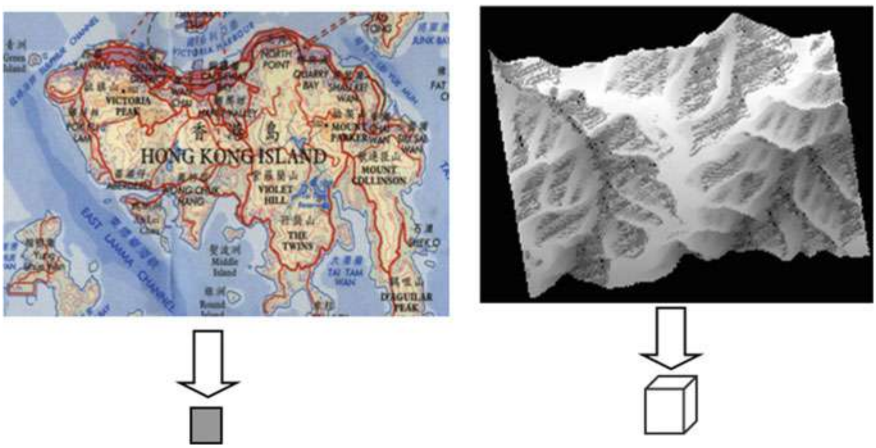
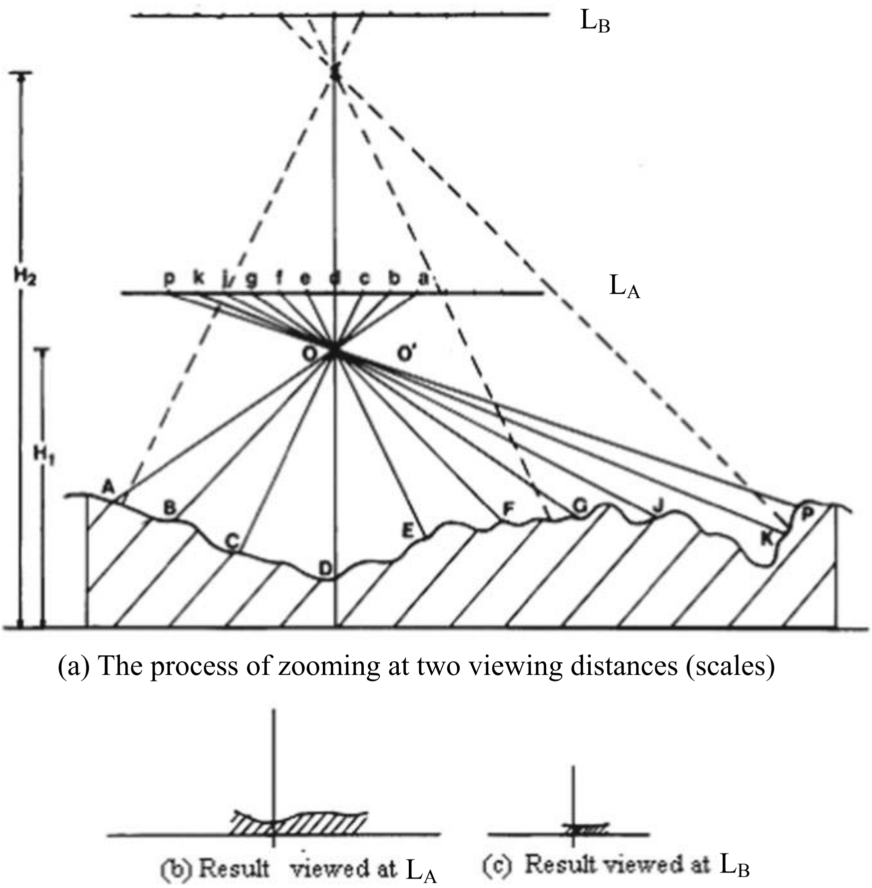


Fig. 8.6 The natural principle: spatial variations within a smallest visible size (SVS) to be neglected (Li 2007)

the idea of this corollary, that is, that all spatial variations within the SVS can be neglected, no matter how big they are on the ground.

Figure 8.7 illustrates the working example of applying the Natural Principle to a terrain surface. Figure 8.7a shows the views of a terrain surface at two different heights based on the Natural Principle, resulting in two quite different representations in terms of complexity. Figure 8.7b, c show the results viewed at levels L_A and L_B, respectively. In these two Figures, the zooming (or generalization) effects are very clear.

To apply the Natural Principle, the critical element to be considered is the value of this "certain limitation" or SVS, beyond which all spatial variations (no matter how complicated) can be neglected. Li and Openshaw (1992, 1993) suggested the following formula:



(a) The process of zooming at two viewing distances (scales)

(b) Result viewed at L_A (c) Result viewed at L_B

Fig. 8.7 Zooming effect of a terrain surface generated by the Natural Principle (Li and Openshaw 1993)

$$K = k \times S_T \times \left(1 - \frac{S_S}{S_T}\right) \tag{8.1}$$

where S_T and S_S are the scale factors of the target and source data, respectively; k is the SVS value in terms of map distance at the target scale and K is the SVS value in terms of ground distance at the target scale. Through intensive experimental testing, Li and Openshaw (1992) recommend a k value between 0.5 and 0.7 mm, i.e.,

$$k = \{0.5\ \text{mm},\ 0.7\ \text{mm}\} \tag{8.2}$$

8.3 Models for Transformations in Scale

To realize a transformation in scale, some transformation models must be adopted and algorithms and/or mathematical functions for these models are applied. The former is the topic of this section and the latter are described in Sect. 8.4.

8.3.1 Data Models for Feature Representation: Space-Primary Versus Feature-Primary

To record features in geographical space, two different viewpoints can be taken: feature-primary and space-primary (Lee et al. 2000).

In a feature-primary view, the geographical space is considered as being tessellated by features and the locations of these features are then determined. This kind of model is also called feature-based. In such a model, features are represented by vectors, leading to the popular term *vector data model*. Figure 8.8a–c show the representation of points, a line and an area using a vector model.

In a space-primary view, the geographical space is considered as being tessellated by space cells. In such a tessellation (partitioning), square raster cells are popularly employed, leading to the popular term *raster data model*. In each raster cell, there could be a feature or there might be no features. A point is represented by a pixel (picture element); a line is represented by a string of connected pixels and an area is formed by a set of connected pixels, as shown in Fig. 8.8d–f. The cells can be in any form, regular or irregular. Irregular triangular networks are another popular tessellation.

On a spherical surface, longitude/latitude is the coordinate system for feature-primary representation. The cells with an equal interval in latitude/longitude (e.g., $6' \times 6'$) are the raster equivalent of spherical tessellation (Fig. 8.9a). However, the actual area size of such a cell varies with the latitude. To overcome this problem, the quaternary triangular mesh (QTM) (Fig. 8.9b) has been used (e.g., Dutton 1984, 1996). The cells can be any shape (e.g., triangle, hexagon), regular or irregular.

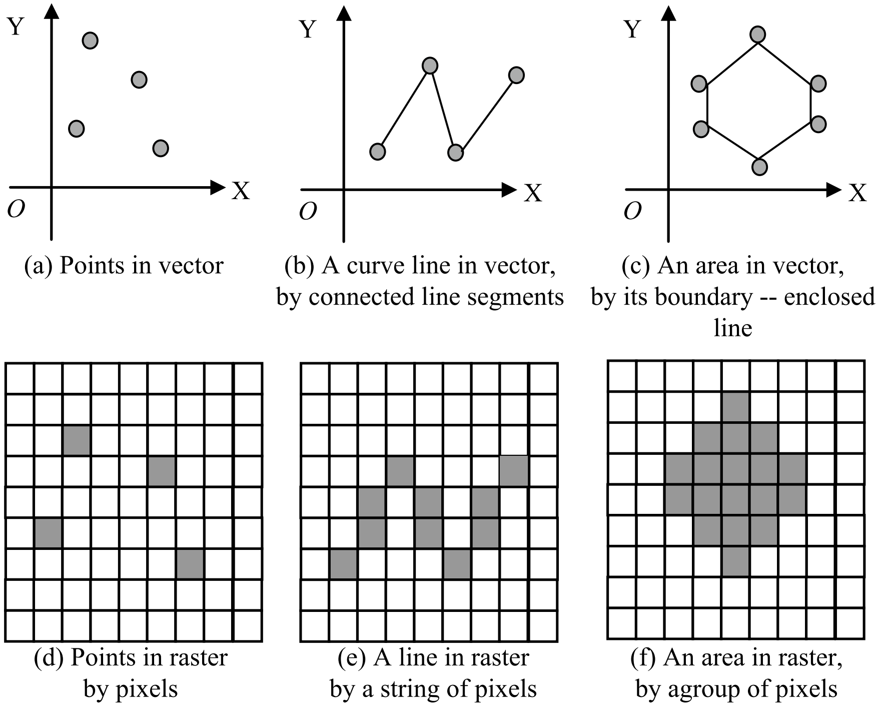


Fig. 8.8 Feature-primary and space-primary representations of spatial features: vector and raster models

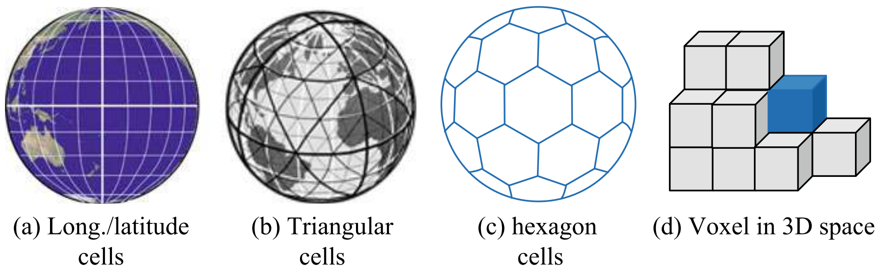


Fig. 8.9 Spatial tessellation of a spherical surface and a 3D space

Figure 8.9c shows the use of a regular hexagon diagram for such a tessellation. For 3D space, the voxel (volume element) is the raster equivalent for space tessellation (Fig. 8.9d).

As the natures of the raster and vector data models are quite different, the model for transformation in scale in these two data models might also differ. Thus, separate subsections are devoted to these topics.

8.3.2 Space-Primary Hierarchical Models for Transformation in Scale

Hierarchical models are popular for the multiscale representation of spatial data at discrete scales. For example, Google Maps, Virtual Earth and Tianditu have all adopted hierarchical models for the representation of images and maps. Figure 8.10 shows the first three zoom levels of the hierarchical model used by Google Maps (Stefanakis 2017). This model has a special name, the pyramid model, which is a result of aggregating a 2 × 2 pixel into one pixel. The number of pixels (squares) at the nth level is 4^{n-1}. A more general form of aggregation is to transform any N × N pixels into one pixel.

A more general form of transformation to create a hierarchical representation is to transform N × N pixels into M × M pixels, e.g., a 5 × 5 into a 2 × 2 or a 3 × 3 into a 2 × 2. In such cases, a resampling process (instead of simple aggregation) is required.

With a hierarchical model, the resolution and cartographic ratio at each level are not necessarily uniform. Typical examples of hierarchical models with variable resolutions are shown in Fig. 8.11, i.e., the quadtree and binary tree models.

With the pyramid and quadtree models, the hierarchical levels are fixed and the transformation in scale jumps from one level to another like stairs. To make the transformation absolutely smooth, we need to make the difference between two steps of the stairs infinitely small, to make the stairs become a continuous linear slope (see Fig. 8.12).

For hierarchical representation on a spherical surface, the Open Geospatial Consortium (OGC) approved a new standard called the Discrete Global Grid System (DGGS) (OGC 2019) The hierarchical representation of QTM as shown in Fig. 8.9b is an example of such a DGGS.

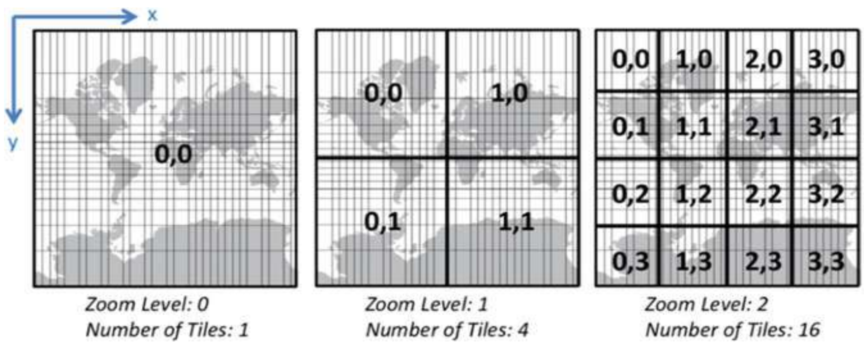


Fig. 8.10 Pyramid model used in Google Maps: the first three zoom levels (Stefanakis 2017)

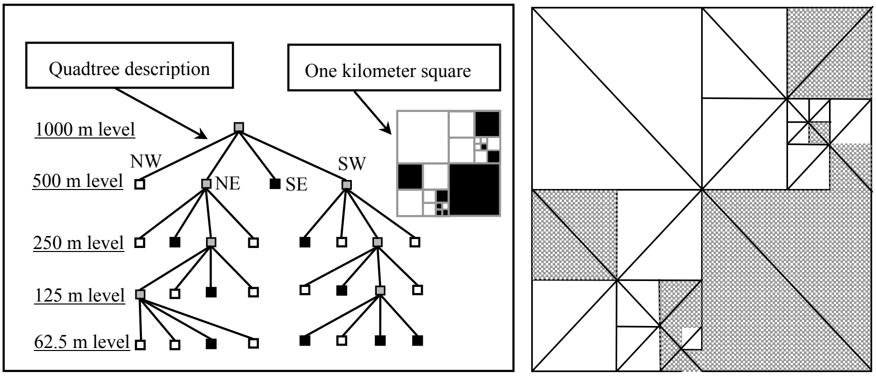


Fig. 8.11 Hierarchical representations of area features with quadtree and binary tree models

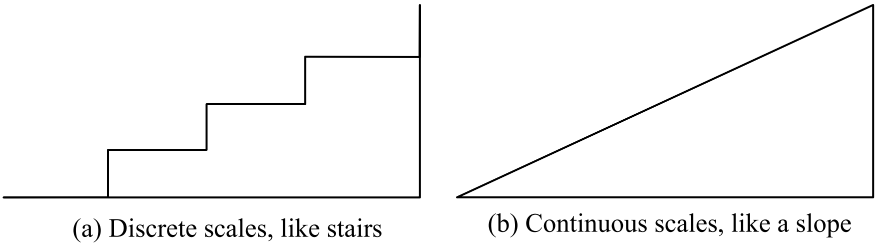


Fig. 8.12 Discrete and continuous transformations in scale: steps and a linear slope

8.3.3 Feature-Primary Hierarchical Models for Transformation in Scale

Hierarchical models have also been used to represent point, line and area features in feature-primary models. Figure 8.13 shows such a representation for the points on a line. At level 1, only two points, i.e., points (1, 1) and (1, 2), will be used to represent the line; at level 2, in addition to the two points at level 1, point 2 will also be used; and at level 3, points (3, 1) and (3, 2) will also be used. This kind of model has been employed for progressive transmission of vector data.

Figure 8.14 shows the hierarchical representation of a river network by the Horton and Shreve models. Figure 8.14a is a hierarchical representation based on river segments. The formation of such a representation starts from the level 1 branches. A segment of level 2 is formed by two or more segments of level 1. Similarly, a segment of level 3 is formed by two or more segments of level 2. All higher level segments are formed by following this principle. Figure 8.14b is a hierarchical representation formed by the Horton model based on a river stroke, which is a concatenated segment. Figure 8.14c is a hierarchical representation formed by the Shreve model. The numbering in this hierarchy is formed by adding the numbers of upstream branches.

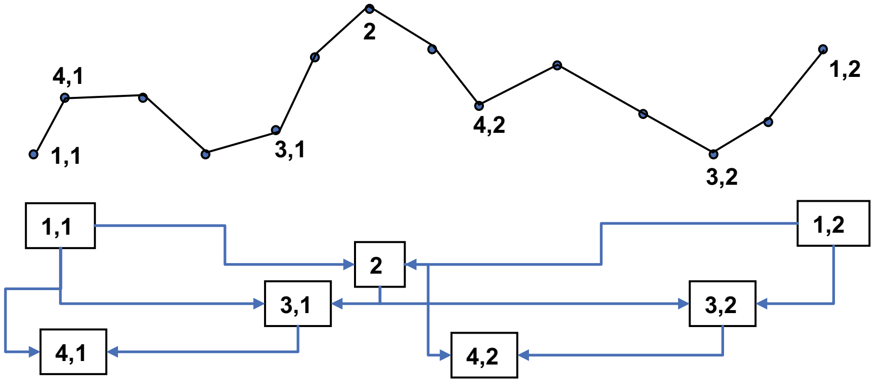
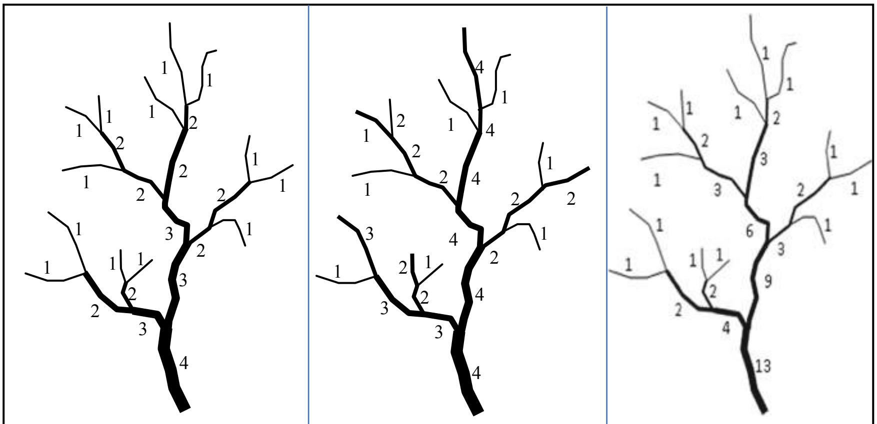


Fig. 8.13 Hierarchical representations of points on a line



(a) Hierarchy formed in downstream routine

(b) Hierarchy formed by the Horton model

(c) Hierarchy formed by the Shreve model

Fig. 8.14 Hierarchical representation of a river network using the Horton and Shreve models (Li 2007)

For example, the ranking value for the segment with the highest ranking is 13, which is a result of adding 9 and 4. Such a numbering of ranking is not continuous.

Figure 8.15 shows the hierarchical representation of two transportation networks. In this case, the importance of each road is evaluated based on geometric information and/or thematic information. A ranking value is assigned to each road.

Figure 8.16 shows a hierarchical representation of area features. The area features in the whole area are first connected by a minimum spanning tress (MST) as a whole group, i.e., Group A. Group A is then subdivided into subgroups B and C by breaking the tree at the connection with the largest span. Similarly, Group B is broken into D and E, and Group C is broken into F and G. The subdivision goes on

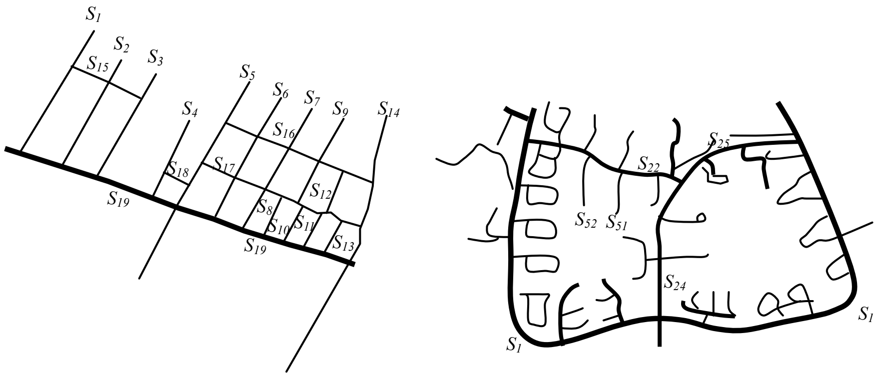


Fig. 8.15 Hierarchical representations of transformation networks (Zhang and Li 2009)

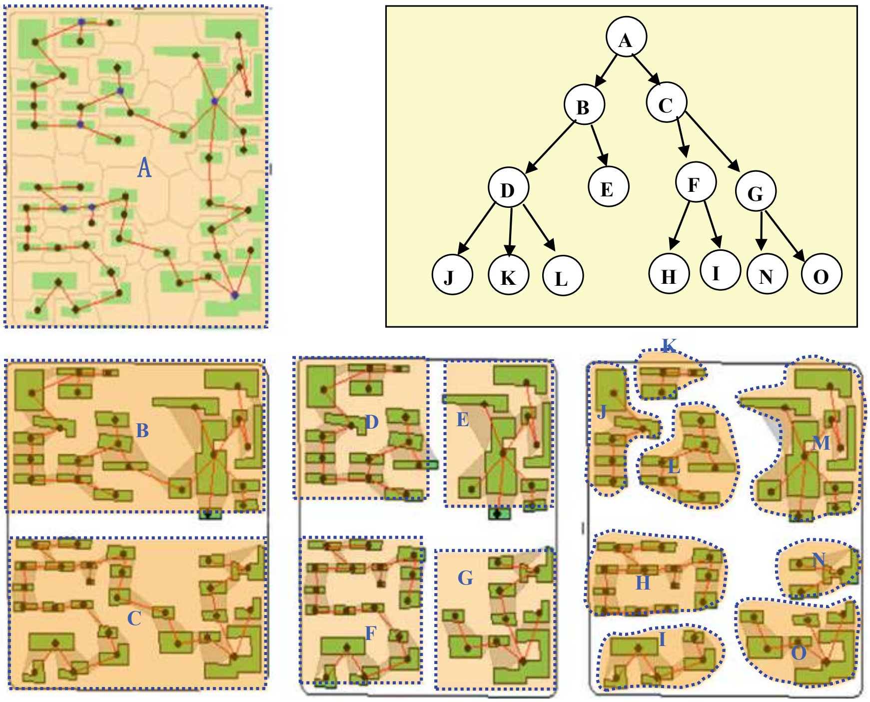


Fig. 8.16 Hierarchical representations of area features (Ai and Guo 2007)

until a criterion is met or until the complete hierarchy is constructed. In the end, a hierarchical representation is formed.

8.3.4 Models of Transformation in Scale for Irregular Triangulation Networks

An irregular triangulation network is an irregular space tessellation that has been widely used for digital terrain models (DTMs). In such a representation, the resolution is variable across the space. Therefore, special models should be used to make the resolution transformable from one to another. Four basic transformation models have been developed for such a purpose (Li 2005):

- *Vertex removal*: A vertex in the triangular network is removed and new triangles are formed.
- *Triangle removal*: A complete triangle with three vertices is removed and new triangles are formed.
- *Edge collapse*: An edge with two vertices is collapsed to a point and new triangles are formed.
- *Triangle collapse*: A complete triangle with three vertices is collapsed to a point and new triangles are formed.

Figure 8.17 illustrates these four transformation models.

8.3.5 Models for Geometric Transformation of Map Data in Scale

The hierarchical model described in Sect. 8.3.2 is suitable to represent raster image data because images are numerical data that naturally record the earth and such a recording follows the Natural Principle described in Sect. 8.2.3. Figure 8.18 shows four images with different resolutions, the result of a "2×2 into "1×1" aggregation. These images appear to be very natural. However, for the categorical data of topographic maps, such a simple transformation does not work well, and there is a need for other transformation models.

Topographic maps are produced via a complicated intellectual process that consists of abstraction, symbolization, generalization, selective omission and simplification. During this process, small details are ignored (or grouped together). All features are represented by symbols (geometric or pictorial). The colors of the symbols are not necessarily the natural colors of features. The graphic symbols are annotated with text (e.g., name of a street/town/city). There are requirements for minimum size, minimum separation and minimum differentiation for graphic elements. Thus, when a map at a larger scale (Fig. 8.19c) is simply reduced by 4 times (equivalent to

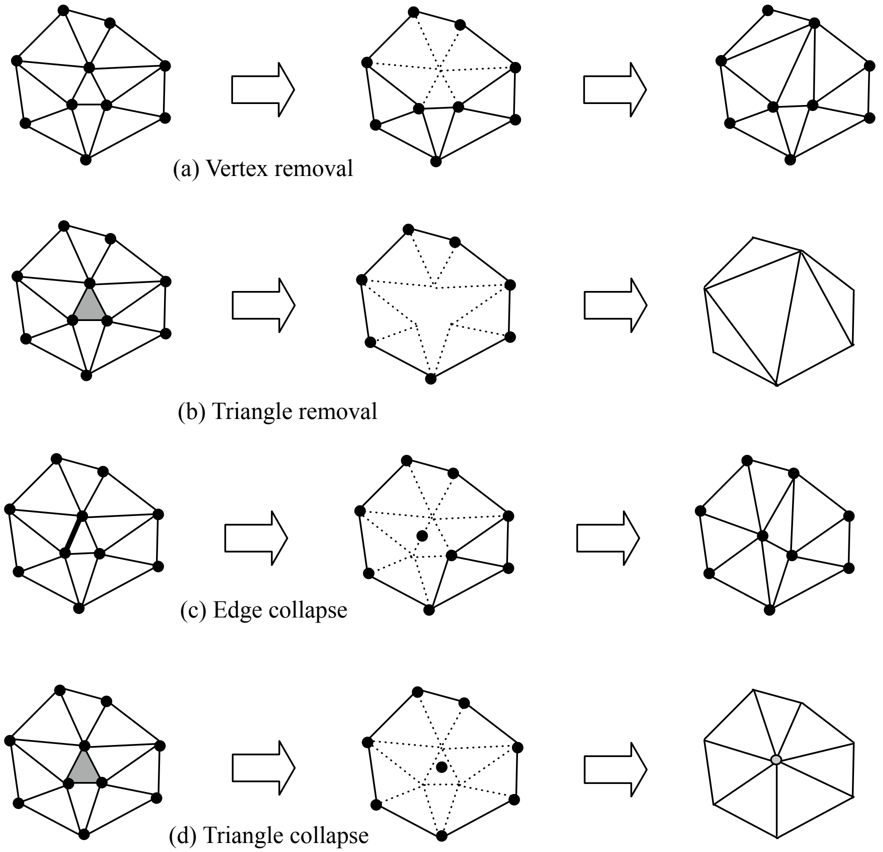


Fig. 8.17 Basic models for geometric transformation in scale for a triangular network (Li 2005)

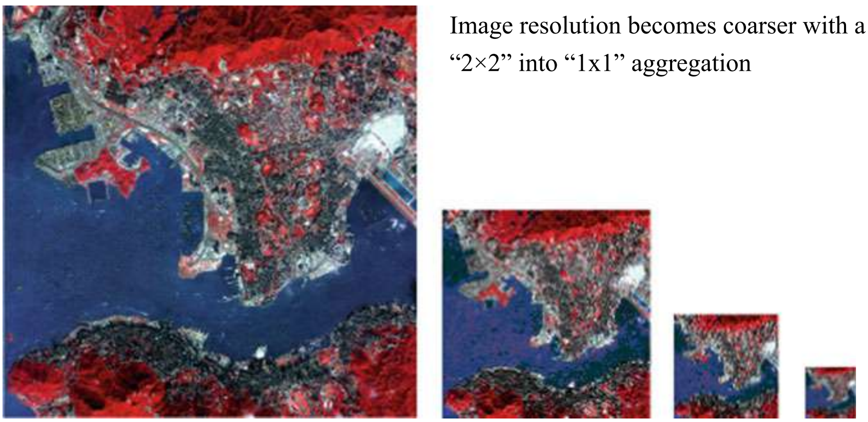
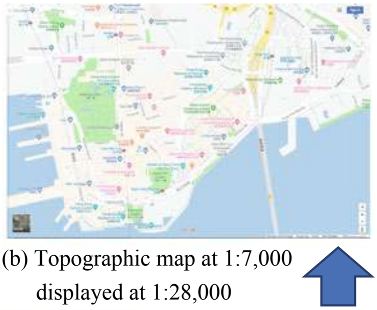
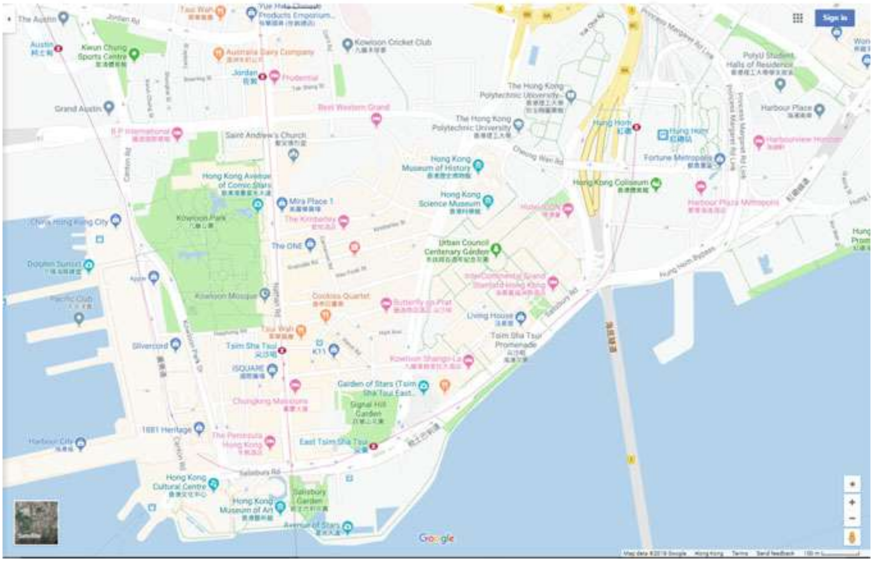


Fig. 8.18 Four images with the same cartographic ratio but different resolutions

(a) Topographic map at 1:28,000



(b) Topographic map at 1:7,000 displayed at 1:28,000



(c) Topographic map at 1:7,000

Fig. 8.19 Kowloon Peninsula represented on maps at two different scales, via generalization and simple scale reduction (extracted from Google Maps)

a "2 × 2 into 1" aggregation), the graphics (Fig. 8.19b) become unclear because the minimum requirements can no longer be met. Figure 8.20 illustrates such a situation with the aggregation of buildings as an example. A set of special models is needed for the transformation of map data from one scale to another to make the graphics at the smaller scale clear (Fig. 8.19a).

The transformation of maps from a larger scale to a smaller scale is called map generalization and has long been studied in the cartographic community. Some transformation models have been identified by researchers. In the traditional textbook by Robinson et al. (1984), only four models are listed, i.e., classification, induction, simplification and symbolization. In the 1980s, more models were identified, and a list of 12 models was produced by McMaster and Shea (1992). Many of these models were still too general to be precisely implemented in a computer system.

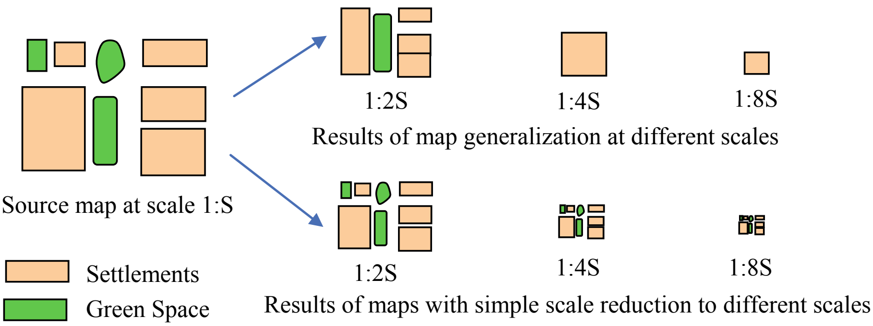


Fig. 8.20 Comparison of map generalization and simple scale reduction

More recently, Li (2007) produced 40 detailed models for implementation. These models are divided into six sets: three sets for individual points, individual lines and individual areas and the other three sets for a class of points, a class of lines and a class of areas. Tables 8.1, 8.2, 8.3, 8.4, 8.5 and 8.6 list the six sets of models.

Table 8.1 Models for geometric transformations in scale of individual point features (Li 2007)

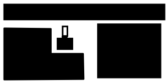
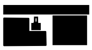
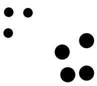
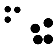
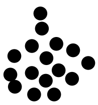
Transformation model	Large-scale	Photo-reduced	Small-scale
Displacement (move because it is too close to another feature)			
Elimination (too small to represent, thus removed)			
Magnification (enlarged due to importance)			

Table 8.2 Models for geometric transformations in scale of a set of point features (Li 2007)

Transformation model	Large-scale	Photo-reduced	Small-scale
Aggregation (group points and make a new one)			
Regionalization (delineate a boundary outlined by points and make a new area feature)			

(continued)

Table 8.2 (continued)

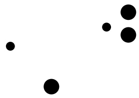
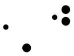
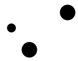
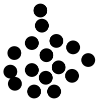
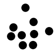
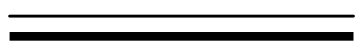
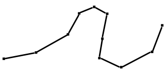
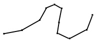
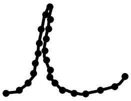
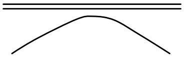
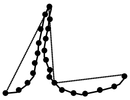
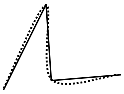
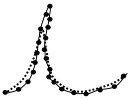
Transformation model	Large-scale	Photo-reduced	Small-scale
Selective Omission (retain more important points and omit less important ones)			
(Structural) Simplification (cluster complexity; the main structure is retained)			
Typification (typical pattern kept while points removed for clarity)			

Table 8.3 Models for geometric transformations in scale of individual line features (Li 2007)

Transformation model		Large-scale	Photo-reduced	Small-scale
Displacement (to move a line away from the position because it is too close to another feature)				
Elimination (to remove the line because it is too minor to be included)				
(Scale-driven) generalization (main structure suitable at target scale retained but small details removed)				
Partial modification (to modify the shape of a segment within a line)				
Point reduction (to reduce the number of points by removing less important points)				
Smoothing (to make the data appear smoother)	*Curve-fitting* (to fit a curve through a set of points)			
	Filtering (to filter out the high-frequency components or small details of a line)			

(continued)

Table 8.3 (continued)

Transformation model	Large-scale	Photo-reduced	Small-scale
Typification (typical patterns of the line bends retained while removing some of them)			

Table 8.4 Models for geometric transformations in scale of a set of line features (Li 2007)

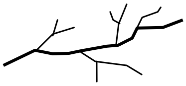
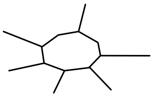
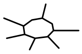
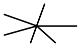
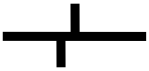
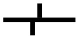
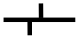
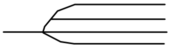
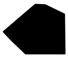
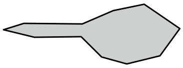
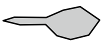
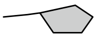
Transformation model		Large-scale	Photo-reduced	Small-scale
Selective omission (to select more important points and remove less important points)				
Collapse (to reduce the dimension)	Ring-to-point			
	Double-to-single			
Enhancement (to keep the characteristics clear)				
Merging (to combine to two or more close lines together)				
Displacement (to move one away from others or both away from each other)				

Table 8.5 Models for geometric transformations in scale of individual area features (Li 2007)

Transformation model		Large-scale	Photo-reduced	Small-scale
Collapse (to reduce the dimension of features)	Area-to-point			
	Area-to-line			
	Partial			
Displacement (to move the area to a slightly different position to solve the conflict problem)				

(continued)

Table 8.5 (continued)

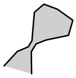
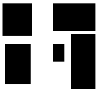
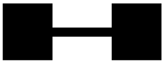
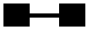
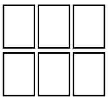
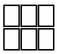
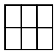
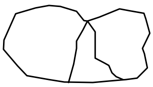
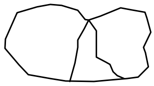
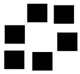
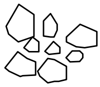
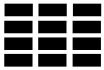
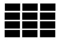
Transformation model		Large-scale	Photo-reduced	Small-scale
Exaggeration (to enlarge one or two dimensions of a small area)	*Directional thickening* (to enlarge an area feature in a direction)			
	Enlargement (to uniformly magnify in all directions)			
	Widening (to widen the bottleneck of an area feature)			
Elimination (to eliminate data that is too small to represent				
(Shape) Simplification (to reduce the complexity of a boundary)				
Split (to split an area into two because the connection between them is too narrow)				

Table 8.6 Models for geometric transformations in scale of a set of area features (Li 2007)

Transformation model	Large-scale	Photo-reduced	Small-scale
Aggregation (to combine area features, e.g., buildings separated by open space)			
Agglomeration (to make area features bounded by thin area features into adjacent area features)			
Amalgamation (to combine area features, e.g., buildings separated by another feature such as roads)			
Dissolving (to split a small area into pieces and merge these pieces into adjacent areas)			

(continued)

Table 8.6 (continued)

Transformation model	Large-scale	Photo-reduced	Small-scale
Merging (to combine two adjacent areas into one)			
Relocation (to move more than one feature around to solve the crowding problem)			
(Structural) Simplification (to retain the structure of area patches by selecting important ones)			
Typification (to retain the typical pattern, e.g., a group of areas aligned in rows and columns)			

8.3.6 Models for Transformation in Scale of 3D City Representations

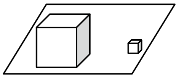
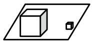
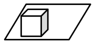
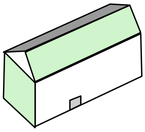
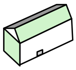
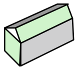
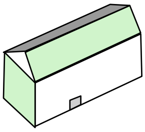
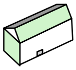
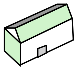
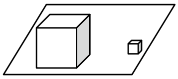
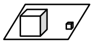
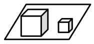
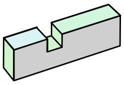
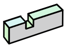
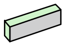
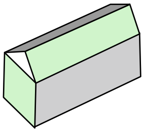
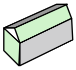
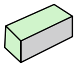
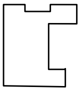
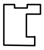
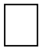
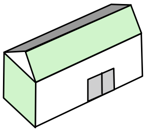
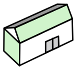
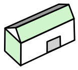
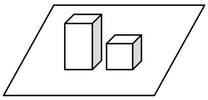
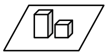
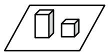
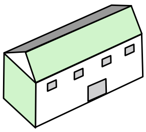
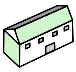
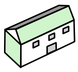
For 3D representation of digital cities, the CityGML, which was officially adopted by the OGC in 2008, specifies five well-defined consecutive levels of detail (LOD) as follows, an example of which is shown in Fig. 8.21 (Kolbe et al. 2008):

- LOD 0—regional, landscape
- LOD 1—city, region
- LOD 2—city districts, projects
- LOD 3—architectural models (outside), landmarks
- LOD 4—architectural models (interior)

For the transformation in scale of 3D features, a set of models is listed in Table 8.7, which is a summary of models proposed in the literature.

Fig. 8.21 The five levels of detail (LoD) defined by CityGML (Kolbe et al. 2008)

Table 8.7 Models for transformation in scale of 3D features

Transformation model		At large scale	Photo-reduced	At small scale
Elimination	Geometric elimination			
	Thematic elimination			
Exaggeration	Thematic exaggeration			
	Geometric exaggeration			
Simplification	Vertical simplification			
	Flattening			
	Squaring			
	Thematic simplification			
Displacement				
Typification				

8.4 Mathematical Solutions for Transformations in Scale

In the previous section, several sets of models for the transformation in scale were described. These models express what is achieved in such transformations, e.g., the shape is simplified, important points retained, and/or the main structure is preserved. To make these transformations work, mathematical solutions (e.g., algorithms and mathematical functions) must be developed for each of these transformations. A selection of these solutions is presented in this section.

8.4.1 Mathematical Solutions for Upscaling Raster Data: Numerical and Categorical

For **raster-based numerical data** such as images and digital terrain models (DTMs), aggregation is widely used to generate hierarchical models. In recent years, wavelet transform (e.g., Mallat 1989), Laplacian transform (Burt and Adelson 1983) and other more advanced mathematical solutions have also been employed. The commonly used aggregation methods are by mode, by median, by average, and by Nth cell (i.e., Nth cell in both the row and column). Figure 8.22 shows a "3 × 3 to 1 × 1" aggregation with these four methods. The 6 × 6 grid is then aggregated into a 2 × 2 grid.

If the new cell interval is not multiples of the original cells, then interpolation must be applied to resample the data. Bilinear and weighted averaging interpolations are widely used for resampling. Figure 8.23 shows the resampling of a 3 × 3 grid into a 2 × 2 grid using weighted averaging interpolation.

Bilinear interpolation can be performed for any four points (not along a line). The mathematical function is as follows:

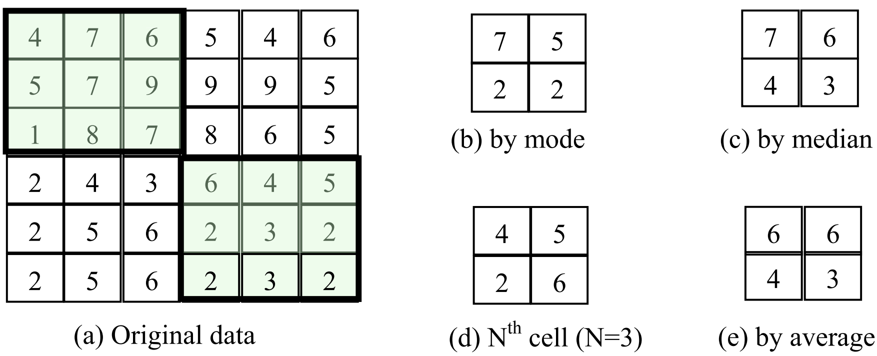


Fig. 8.22 "3 × 3 to 1 × 1" aggregation of numerical data

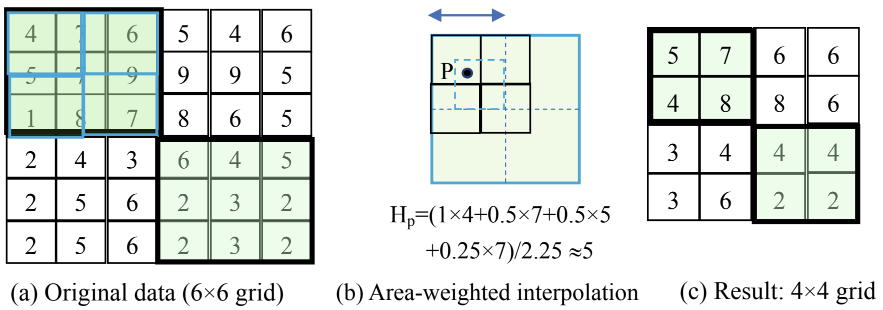


Fig. 8.23 "3 × 3 to 2 × 2" resampling of numerical data

$$z = a_0 + a_1 x + a_2 y + a_3 xy \tag{8.3}$$

where a_0, a_1, a_2, a_3 is the set of four coefficients, which are to be determined by four equations that are formed by making use of the coordinates of four reference points, i.e., the centers of the four grid cells in Fig. 8.23b: $P_1(x_1, y_1, z_1)$, $P_2(x_2, y_2, z_2)$, $P_3(x_3, y_3, z_3)$ and $P_4(x_4, y_4, z_4)$. The mathematical formula is as follows:

$$\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 & x_1 y_1 \\ 1 & x_2 & y_2 & x_2 y_2 \\ 1 & x_3 & y_3 & x_3 y_3 \\ 1 & x_4 & y_4 & x_4 y_4 \end{bmatrix}^{-1} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} \tag{8.4}$$

Once the coefficients a_0, a_1, a_2, a_3 are computed, the height Z_P of any point P with a given set of coordinates (x_P, y_{P}) can be obtained by substituting (x_P, y_{P}) into Eq. (8.1).

The mathematical expression of weighted averaging interpolation is as follows:

$$z = \frac{\sum_{i=1}^{n} w_i z_i}{\sum_{i=1}^{n} w_i} \tag{8.5}$$

where w_i is the weight of the ith reference point; z_i is the height of the ith reference point; and n is the total number of the reference points used. In the case of Fig. 8.23b, n = 4.

Weights may be determined by using different functions. The simplest weighting function assigns an equal weight to all reference points. However, it seems unfair to those reference points that are closer to the interpolation point, as such points should have a higher influence on the estimate. As a result, distance-based or area-based weighting are more commonly used. The inverse of distance is most popularly used:

$$w = \frac{1}{d} \text{ or } w = \frac{1}{d^2} \tag{8.6}$$

where d is the distance from a reference point to the interpolation point. In the case of interpolating the height of P in Fig. 8.23b, the four distances from the four (old) cell centers to point P will be used. Figure 8.23b also shows that the distance of each cell center to the interpolation point P is directly related to the size of the area contributed by each (old) cell to the new cell. If the area size is denoted as A, the weighting function is

$$w_i = A_i \tag{8.7}$$

For example, if the area of the new cell is composed of 100% of the upper left cell, 50% of the upper right cell, 50% of the lower left cell and 25% of the lower right cell, the weights of these four cells are 1.0, 0.5, 0.5 and 0.25, and the result of the interpolation is:

$$z_p = 1 \times 4 + 0.5 \times 7 + 0.5 \times 5 + 0.25 \times 7)/2.25 \approx 5$$

For the **raster-based categorical data**, the averaging and median are no longer applicable. The mode (also called the majority in some literature) is still valid and widely used. Figure 8.24b shows such a result. However, the value for the upper right cell is difficult to determine as there is no mode (majority) in the 3 × 3 window at the upper right corner of the original data (Fig. 8.24a). Notably, some priority rules or orders are in practical use. For example, a river feature is usually given a priority because thin rivers are likely to be broken after aggregation. Figure 8.25 shows the improvement in the connectivity of river pixels with water as the priority. Figure 8.24c-e show the results with different options, e.g., random selection and central pixel. It is also possible to consider the statistical distribution of the original data (e.g., A = 8, T = 10, W = 6, S = 11) to try to maintain the distribution as much as possible.

In the aggregation/resampling process, as illustrated in Figs. 8.22, 8.23 and 8.24, a moving window is used but the question of the most appropriate window size has rarely been addressed. Li and Li (1999) suggested that the size of the moving window for aggregation/resampling should be computed based on the resolutions

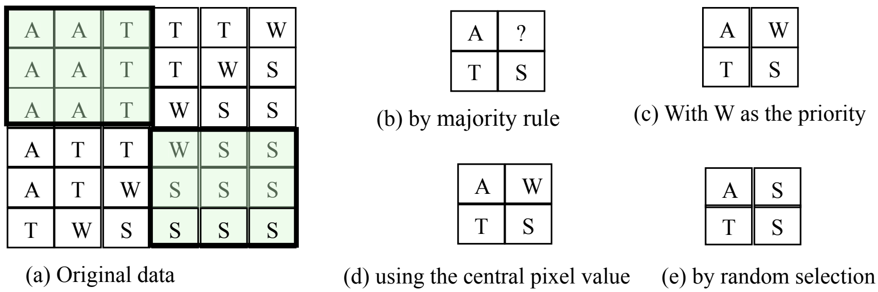
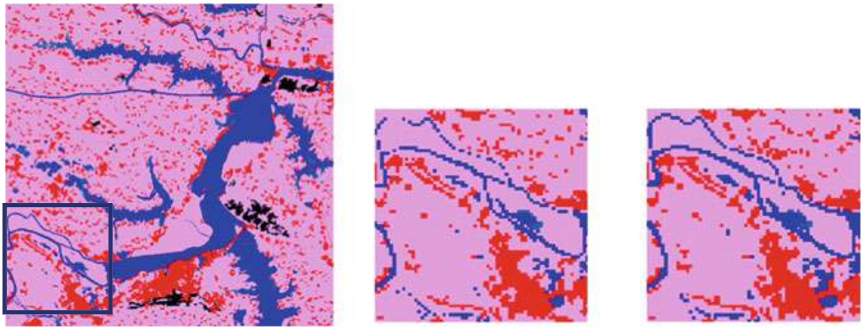


Fig. 8.24 Aggregation of raster-based categorical data



(a) Original data (b) “3×3 to 1” aggregation (c) “3×3 to 1” aggregation with water as the priority

Fig. 8.25 Aggregation of landcover data with priority (extracted from Tan 2018)

(scales) of the input and output, following the *Natural Principle* (Li and Openshaw 1993) described in Sect. 8.2.3. Mathematically,

$$W = \frac{K}{R_{in}} \tag{8.8}$$

where R_{in} is the resolution (scale) of the input data; K is the SVS value in terms of ground distance at the target scale computed by Eq. (8.1), and W is the size of the window’s side in terms of pixel numbers (of input data).

8.4.2 Mathematical Solutions for Downscaling Raster Data

Downscaling produces a finer spatial resolution raster data than that of the input data through prediction. It is possible to use simple resampling (as described in Sect. 8.4.1) to achieve downscaling. However, methods based on spatial statistical analysis are more theoretically grounded and have become popular (Atkinson 2008, 2013), particularly area-to-point prediction (ATPP). Double dictionary learning has also been used (Xu and Huang 2014).

Area-to-point kriging (ATP Kriging or ATPK) (Kyriakidis 2004) is the typical method. ATP Kriging can ensure the coherence of predictions, such as by ensuring that the sum of the downscaled predictions within any given area are equal to the original aggregated count. Some variants of ATP Kriging have also been developed, e.g., ATP Poisson Kriging (Goovaerts 2008, 2009, 2010), indicator cokriging (Boucher and Kyriakidis 2006) and ATP regression Kriging (Wang et al. 2015). In this section, the base version of ATP Kriging is described.

The basic principle behind Kriging is weighted averaging. The weights are optimized by using the semivariogram computed from the original data.

$$Z_{e,p} = \sum w_i \times Z_i \tag{8.9}$$

where $Z_{e,p}$ is the estimated (interpolated) value; Z_i is the value of the ith reference point; w_i is the value of the ith reference point and $\sum w_i = 1$.

The interpolated value $Z_{e,p}$ is very likely to deviate from the actual value at point p, $Z_{a,p}$. The difference is called the estimation error. The variance of these deviations is expressed by Eq. (8.10).

$$\sigma_z^2 = \frac{\sum_{i=0}^{n} (Z_{e,p} - Z_{a,p})_i^2}{n} \tag{8.10}$$

The basic principle of Kriging is to produce the minimum estimation variance by choosing a set of optimal weights. Such weights are obtained by solving a set of simultaneous equations:

$$\begin{aligned} w_1 \times \gamma(d_{11}) + w_2 \times \gamma(d_{12}) + \cdots\cdots + w_m \times \gamma(d_{1m}) + \lambda &= \gamma(d_{1P}) \\ w_1 \times \gamma(d_{21}) + w_2 \times \gamma(d_{22}) + \cdots\cdots + w_m \times \gamma(d_{2m}) + \lambda &= \gamma\left(d_{2p}\right) \\ \cdots\cdots \\ w_{1m} \times \gamma(d_{m1}) + w_{12} \times \gamma(d_{m2}) + \cdots\cdots + w_{1m} \times \gamma(d_{mm}) + \lambda &= \gamma(d_{mP}) \\ w_1 + w_2 + \cdots\cdots + w_m &= 1 \end{aligned} \tag{8.11}$$

where w_i is the weight of the ith reference point; λ is the Lagrange multiplier; and $\gamma(d)$ is the semivariogram value of points with distance d apart, which can be expressed as follows:

$$\gamma(d) = \frac{\sum_{i=0}^{n_d} (Z_i - Z_{i+d})_i^2}{n_d} \tag{8.12}$$

In ATP Kriging, the interpolation finds an estimate for a point at higher resolution. In such a case, a cell point at coarser resolution corresponds to an area at higher resolution. Therefore, the set of simultaneous equations is as follows:

$$\begin{aligned} w_1 \times \gamma(d_{11}) + w_2 \times \gamma(d_{12}) + \cdots\cdots + w_m \times \gamma(d_{1m}) + \lambda &= \gamma(d_{1A}) \\ w_1 \times \gamma(d_{21}) + w_2 \times \gamma(d_{22}) + \cdots\cdots + w_m \times \gamma(d_{2m}) + \lambda &= \gamma(d_{2A}) \\ \cdots\cdots \\ w_{1m} \times \gamma(d_{m1}) + w_{12} \times \gamma(d_{m2}) + \cdots\cdots + w_{1m} \times \gamma(d_{mm}) + \lambda &= \gamma(d_{mA}) \\ w_1 + w_2 + \cdots\cdots + w_m &= 1 \end{aligned} \tag{8.13}$$

where $\gamma(d_{iA})$ is the point-to-block semivariogram value from the ith point to area A. It is the same as the average of the point-to-point semivariogram value between the ith point and the points within A.

8.4.3 Mathematical Solutions for Transformation (in Scale) of Point Set Data

As discussed in Sect. 8.3.5, a number of transformations are possible, such as regionalization, aggregation, selective omission, structural simplification, and typification. In both aggregation and regionalization, the clustering plays a central role. In aggregation, a cluster is represented by a point; in regionalization, a cluster is represented by an area. Thus, clustering is discussed here.

Clustering is one of the most primitive activities of human beings (Anderberg 1973; Xu and Wunsch 2005). Clustering of spatial points is one of the main tasks in digital earth such as in spatial data mining and exploratory spatial analysis (Estivill-Castro and Lee 2002; Miller and Han 2009; Openshaw et al. 1987). Numerous clustering methods are available. The classic algorithms are the K-means algorithms, and the ISODATA algorithm is an important extension of K-means (Ball and Hall 1967). Classification by K-means is achieved by minimizing the sum of the square error over all K clusters (i.e., the objective function) as follows:

$$E = \sum_{k=1}^{K} \sum_{x_i \in C_k} \left| x_i - \bar{C}_k \right|^2 \tag{8.14}$$

where $\bar{C}_k$ is the mean of the cluster C_k. The procedure of this algorithm is as follows:

(1) arbitrarily select K points from data set (X) as initial cluster centroids;
(2) assign each point in X to the cluster whose centroid is closest to the point;
(3) compute the new cluster centroid for each cluster; and
(4) repeat Steps (2) and (3) until no change can be made.

However, Li et al. (2017) noted that (a) all clustering algorithms discover clusters in a geographical dataset even if the dataset has no natural cluster structure and (b) quite different results will be obtained with different sets of parameters for the same algorithm. These two problems lead to the difficulty in understanding the implications of the clustering results. Consequently, Li et al. (2017) proposed a scale-driven clustering theory. In this theory, scale is modeled as a parameter of a clustering model; the scale dependency in the spatial clustering is handled by constructing a hypothesis testing; and multiscale significant clusters can be discovered by controlling the scale parameters in an objective manner. The basic model can be written as

$$\mathrm{C} = f(D, A) \tag{8.15}$$

where C is the clustering result; f is the clustering model; A is the analysis scale (the size of clusters or the degree of homogeneity within clusters); and D is the data scale (e.g., resolution and extent).

The clustering consists of two major tasks, i.e., estimation of the density for each point and detection of dense regions. The procedure is as follows:

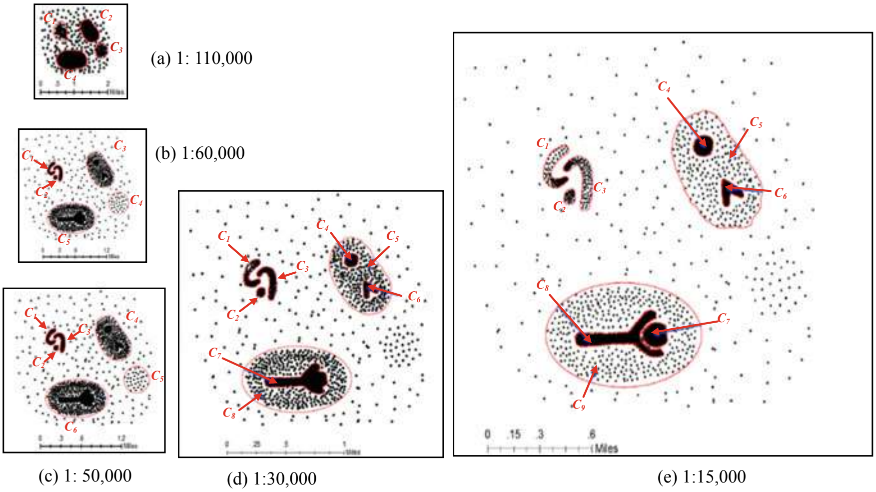


Fig. 8.26 Scale-driven clustering: five results produced at five different scales from the same simulated dataset (Li et al. 2017)

(1) Control the data scale: Determine the SVS (smallest visible size) based on input and output data scales and following the Natural Principle, and ignore all the points within an SVS in the calculation of point data density.
(2) Identify high-density points: The probability density function (PDF) of the dataset is estimated with adaptive analysis scales. The PDF are statistically tested against a null distribution. Points with a significantly higher density are then identified.
(3) Group the high-density points into clusters: Clusters with different densities are formed by adaptively breaking the long edges in the triangulation of high-density points. The significance of clusters obtained at multiscales can be statistically evaluated.

Figure 8.26 shows an example of transforming a set of point data into five different scales. When the output scale decreases (or the resolution becomes coarser), fewer classes can be identified by this clustering technique.

8.4.4 Mathematical Solution for Transformation (in Scale) of Individual Lines

As discussed in Sect. 8.3.5, there are eight different types of transformation for individual lines and the algorithms/mathematical solutions for the transformation models are discussed in detail by Li (2007). In this section, two classic algorithms are described in detail, i.e., the Douglas–Peucker algorithm (Douglas and Peucker 1973) and the Li–Openshaw algorithm (Li and Openshaw 1992).

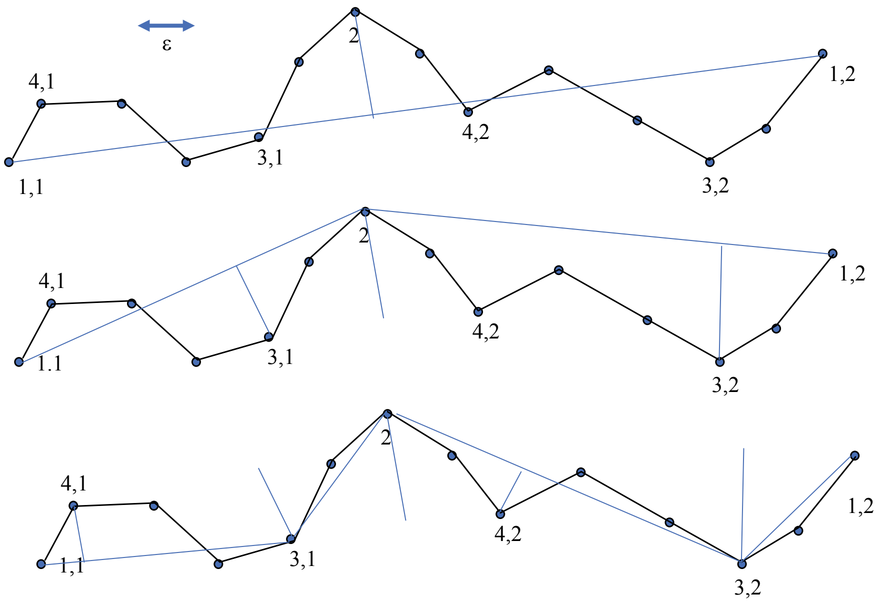


Fig. 8.27 Douglas–Peucker algorithm for generation of a point hierarchy

In Fig. 8.13, a hierarchical representation of the points on a line is presented. The order of these points is sorted by the Douglas–Peucker algorithm. The working principle of this algorithm is illustrated in Fig. 8.27. A curve line is given with an ordered set of points, and a distance tolerance ε (> 0) is set. The basic idea is to use a straight line connecting the first and last points to represent the curve line if the deviations from all line points to the straight line are smaller than ε. In this case, only the two end points are selected and all middle points are regarded as being insignificant and can be removed.

The algorithm first selects two end points (i.e., the first and last points). It then searches for the point that has the largest deviation from the straight-line segment connecting these two end points, i.e., at point 2 in Fig. 8.27. If the deviation is larger than ε, then this point is selected; otherwise, all other points can be ignored. In this example, point 2 is selected and it splits the line into two pieces. The search is then carried out for both pieces. Then, points (3, 1) and (3, 2) are selected. These two points split the whole line into four pieces, and the search will be carried out for these four pieces. The process continues until all the deviations are smaller than ε.

Visvalingham and Whyatt (1993) and Li (2007) noted that the Douglas–Peucker algorithm may cause huge shape distortion. To overcome this problem, Visvalingham and Whyatt (1993) believed that the size of an area "sets a perceptual limit on the significance" and is the most reliable metric for measuring the importance of points since it simultaneously considers the distance between points and angular measures. They used the *effective area* of a point as the threshold, as illustrated in Fig. 8.28. For example, the effective area of point 2 is the area covered by the triangle formed

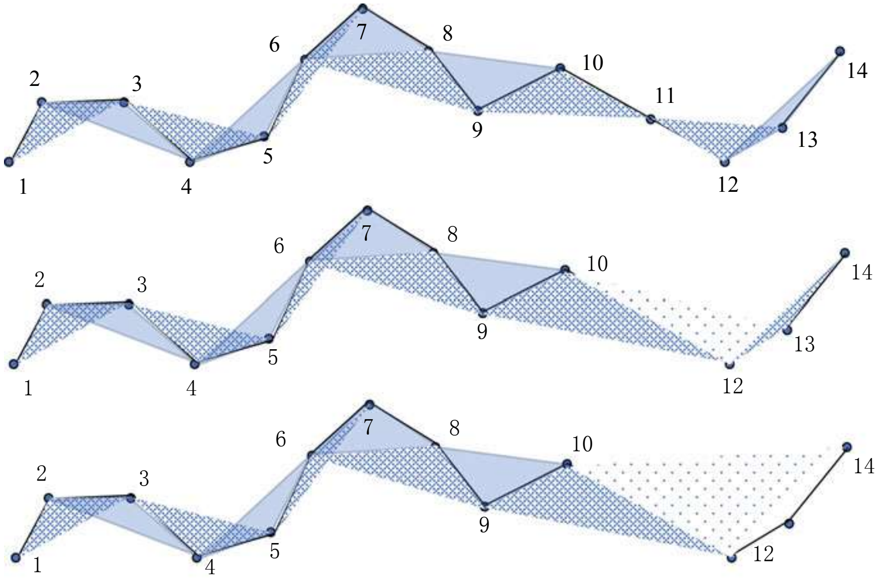
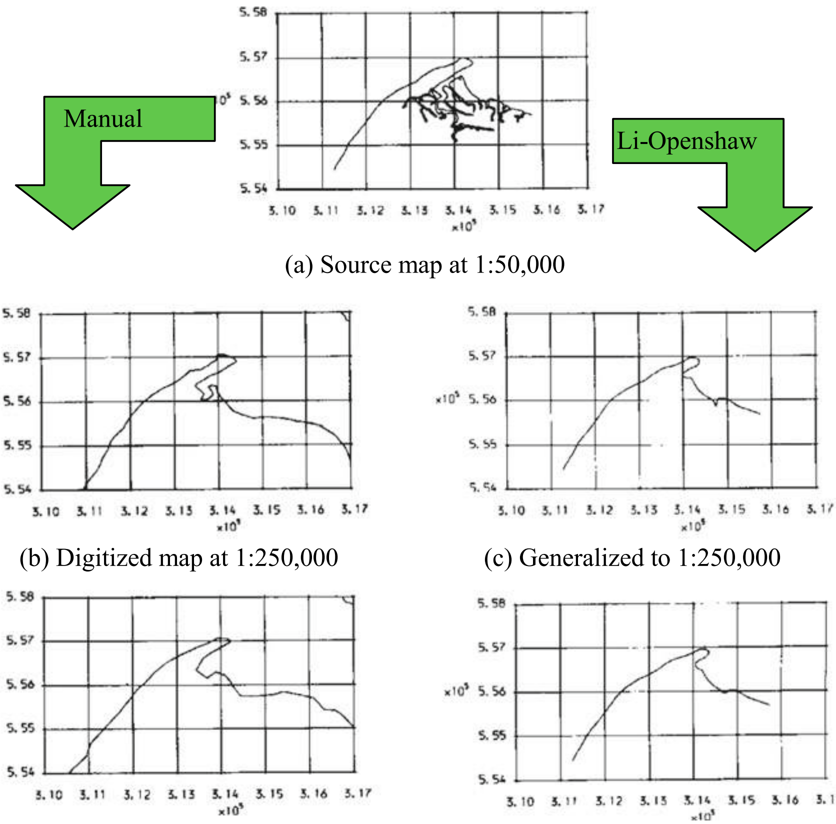


Fig. 8.28 Effective area as a metric in Visvalingham–Whyatt algorithm for generation of a point hierarchy

by points 1, 2 and 3. The basic idea of this algorithm is to progressively eliminate the point with smallest effective area from the list, and the effective areas of the two points adjoining the recently deleted point should be immediately updated. In this example, point 11 is first eliminated and point 13 is removed. The points are ranked from least to most important according to the sequence of elimination.

Many researchers (Li and Openshaw 1992; Visvalingham and Whyatt 1993; Weibel 1996) have noted that the Douglas–Peucker algorithm will create self-intersection (with the line itself) and cross-intersections (between neighboring lines). This problem is associated with all the algorithms with an objective of point reduction or curve approximation. Li and Openshaw (1992) argued that these algorithms are not suitable for generalization (i.e., transformation in scale) because they are normally evaluated with the original curve line (but do not correspond with the curve line at other scales) as the benchmark. To perform transformation in scale for line features, the Li–Openshaw algorithm should be employed as this algorithm, "by virtue of its raster structure, implicitly (but not explicitly) avoids self-overlaps" (Weibel 1996). Even for a very complex coastline, it can produce results that are extremely similar to those manually generalized to various scales, as illustrated by Fig. 8.29. Many recent evaluations also indicate that the Li–Openshaw algorithm produces reasonable and genuine results (e.g., Zhu et al. 2007).

The Li–Openshaw algorithm follows the *Natural Principle* (Li and Openshaw 1993) described in Sect. 8.2.3, i.e., to neglect all spatial variations within the SVS that is computed by using input and output scales. The SVS is mimicked by a cell or



(a) Source map at 1:50,000

(b) Digitized map at 1:250,000

(c) Generalized to 1:250,000

(d) Digitized map at 1:625,000

(e) Generalized to 1:500,000

Fig. 8.29 A comparison of the results of manual generalization and the Li–Openshaw algorithm (Li 2007)

pixel although other geometric elements are also possible (e.g., hexagon by Raposo in 2013). The cells can be organized in the form of a none overlapped tessellation or with overlaps. If there is no overlap, it becomes a pure raster template. Figure 8.30 shows the generalization (transformation) process with a raster template. In this example, each SVS is represented by a raster pixel and the result is represented by pixels, as shown in Fig. 8.30b, or by its geometric center.

Three algorithms were developed by Li and Openshaw (1993) in different modes, raster node, vector mode and raster-vector mode. The algorithm in raster-vector mode was recommended. Figure 8.31 shows the generalization by the Li–Openshaw algorithm in raster-vector mode. The first point to be recorded is the starting point. The second point is somewhere within the second cell. In this implementation, the middle point between the two intersections between cell grids and the line (Fig. 8.31b) is used. If there is more than one intersection, the first (from the inlet direction) and

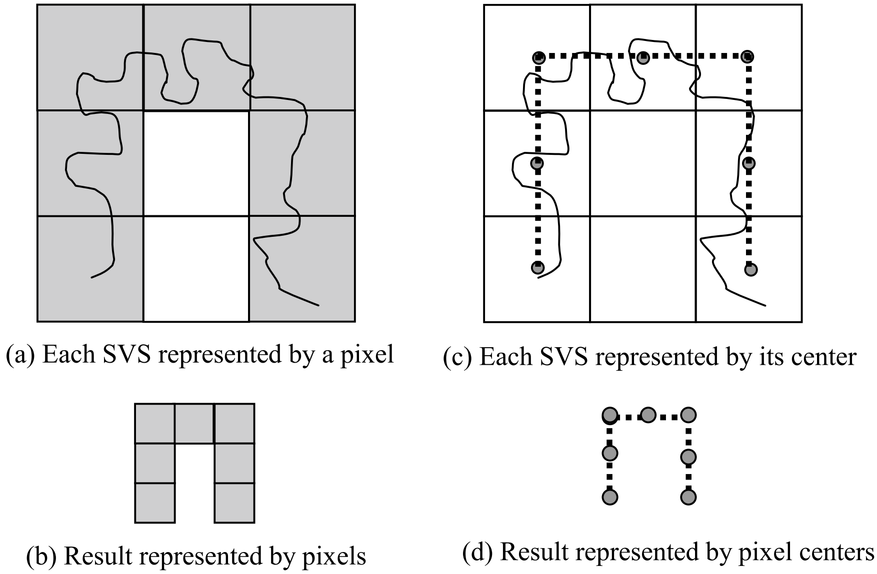


Fig. 8.30 Li–Openshaw algorithm in raster mode; each cell is an SVS (Li 2007)

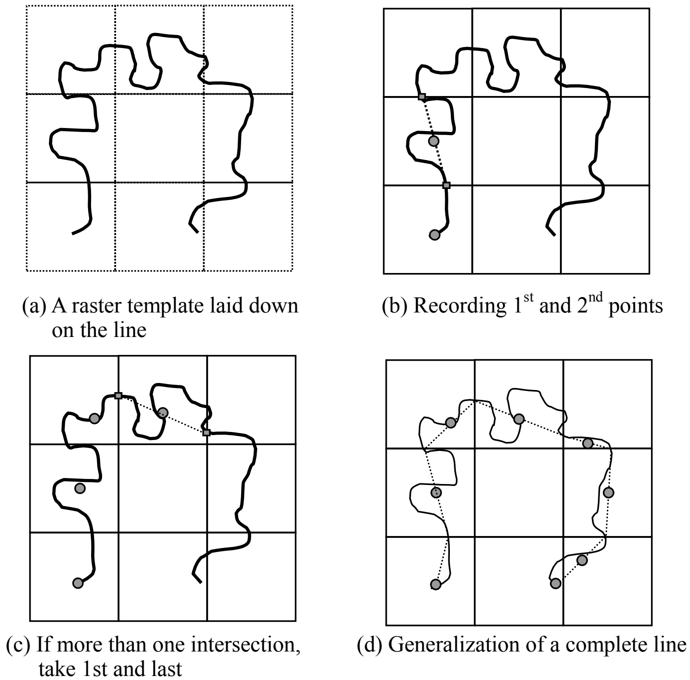


Fig. 8.31 Li–Openshaw algorithm in raster-vector mode (Li 2007)

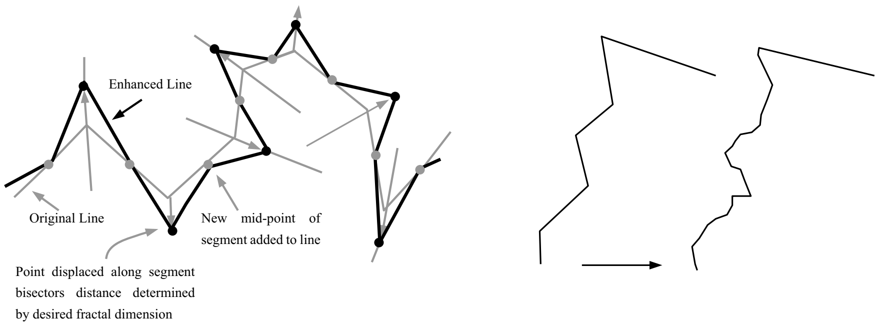


Fig. 8.32 Downscaling of a line by fractal enhancement (Clarke 1995)

the last (outlet direction) intersections are used to determine the position of the new point ((Fig. 8.31c). The final result of the generalization of a complete line is given in Fig. 8.31d.

Similar to the algorithm in raster mode, overlap between SVSs can also be adopted, although it is not too critical. Notably, it is not necessary to take the average to represent a cell. It does not matter what point within the cell is used, as the cell itself is an SVS. Thus, it is also possible to take an original point, which is considered a critical point to represent the cell.

Some work has also been carried out to downscale the lines, i.e., to add more details to the lines. A typical example of such work is that by Dutton (1981), which adds more details to the line by following the fractal characteristics of the line itself (see Fig. 8.32).

8.4.5 Mathematical Solutions for Transformation (in Scale) of Line Networks

In geographical space, three types of line networks are commonly used, contour line networks, hydrological networks and transportation networks. Some hierarchical models were presented in Sect. 8.3.3. The mathematical solutions for the transformation in scale of these networks are discussed in detail by Li (2007). Here, only the construction of a hierarchy for transportation networks is described.

The first approach is based on the *importance of roads*. As road networks are stored in segments and intersections in a database, two steps are required, to build strokes and to order strokes, as illustrated in Fig. 8.33. To build strokes means to concatenate continuous and smooth network segments (see Fig. 8.33a) into a whole (see Fig. 8.33b). To order strokes means to rank the strokes in a descending order based on their importance from high to low (see Fig. 8.33b). The importance of each stroke can be calculated according to various properties, i.e., geometric properties such as length (Chaudhry and Mackaness 2005), topological properties such

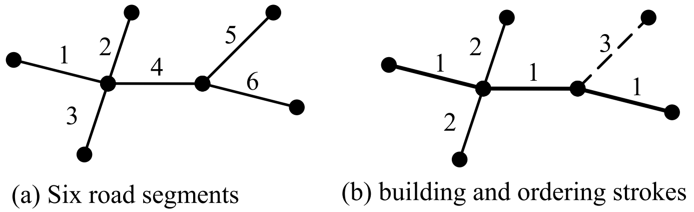


Fig. 8.33 Stroke formation and ordering

as degree, closeness and/or betweenness (Jiang and Claramunt 2004), and thematic properties such as road class. A comparative analysis of the methodology for building strokes was carried by Zhou and Li (2012). With each stroke, given an importance, a stroke-based hierarchy of a line network can be built.

The importance of strokes can be evaluated by the connectivity of strokes in the network. ego-network analysis and weighted ego-network analysis are possible methods (Zhang and Li 2011). Figure 8.34 shows the basic structure of three types of ego-networks and the weight of each link, also called the proportional link strength.

The proportional link strength of each link (p_{ij}) from node i to any of its immediate neighbor nodes can be defined as the reciprocal of the degree of connectivity (k) of node i. Mathematically,

$$p_{ij} = \frac{1}{k_i} \qquad (j \in i_{ne}) \tag{8.16}$$

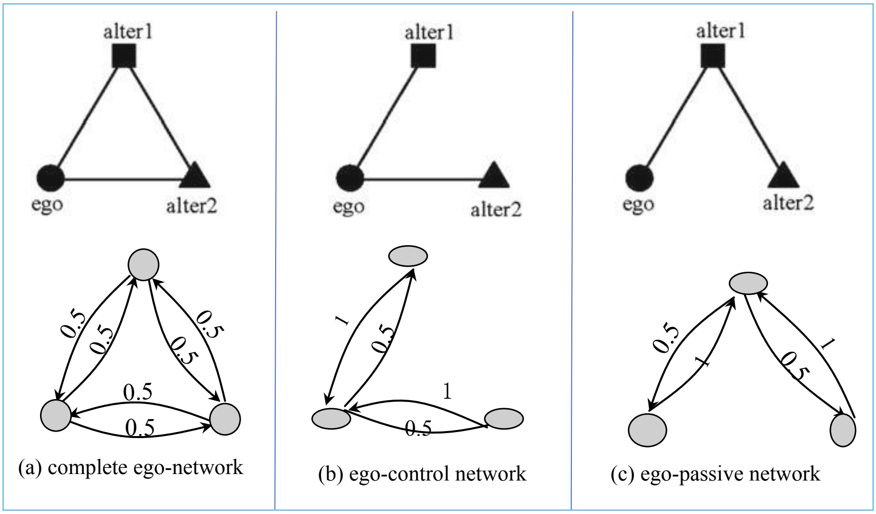


Fig. 8.34 Ego-networks and proportional link strength

For instance, in Fig. 8.34a, the ego is connected to both alter1 and alter2, so its degree of connectivity is 2; thus, the strengths of links from this ego to alter1 and to alter2 are both 1/2 = 0.5. The strengths of other links are also indicated in Fig. 8.34.

If node i and node j are not directly linked but are linked via another node q in the neighbor (*ne*), the strength of the link from node i to node j *(i.e.,* p_{ij}) is defined as:

$$p'_{ij} = p_{iq} p_{qj} \tag{8.17}$$

The total link strength (C_{ij}) from node i to node j is defined as the square of the sum of the direct link strength and the indirect link strength from node i to node j. Mathematically,

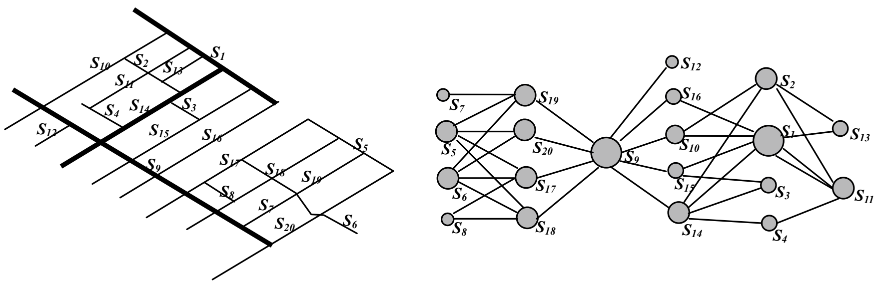
$$C_{ij} = \left(p_{ij} + \sum p'_{ij}\right)^2 = \left(p_{ij} + \sum_{q=1}^{m} p_{iq} p_{qj}\right)^2 \tag{8.18}$$

The C_{ij} value reveals the constraint of i by j. The larger the C value is, the larger the constraint over i, and the smaller the opportunity for i.

To apply this concept to a transport network, the physical road network is first concerted into a connectivity graph, and the link strength values are computed for each node in the connectivity graph. Figure 8.35 shows an example. Roads can then be ranked by the link strength values.

The ego-network is a feasible and effective solution for the formation of hierarchies for road networks. However, Zhang and Li (2011) identified two significant limitations, the deviation of the link intensity definition from reality and the so-called 'degree 1 effect'. They subsequently developed a weighted ego-network analysis method.

Another important development is the *mesh density-based approach* proposed by Chen et al. (2009). The so-called mesh is a closed region surrounded by several road segments. In this approach, the density of each mesh in the road network is computed according to the following formula:



(a) A regular road network (b) The connectivity graph

Fig. 8.35 Formation of a network hierarchy by ego-network analysis (Zhang and Li 2009)

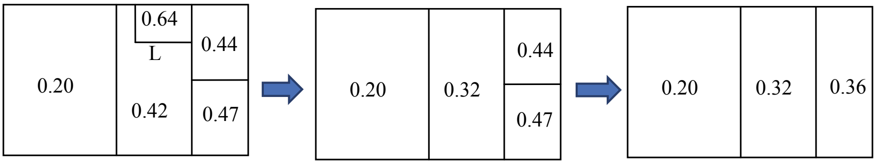


Fig. 8.36 Mesh density-based approach

$$D = \frac{P}{A} \tag{8.19}$$

where P is the perimeter of the mesh and A is the area of the mesh.

Then, the meshes with the highest density are merged progressively, as illustrated in Fig. 8.36. In this Figure, the mesh with density of 0.64 is first merged into the that with a density of 0.42 and segment L is eliminated. The density (0.32) of the new mesh is then updated. The process is iterated until only one mesh is left.

Generally, a road network is often a hybrid of linear and areal patterns, thus Li and Zhou (2012) proposed the construction of hybrid hierarchies, i.e., an integration of a line hierarchy and an area hierarchy.

8.4.6 *Mathematical Solutions for Transformation of a Class of Area Features*

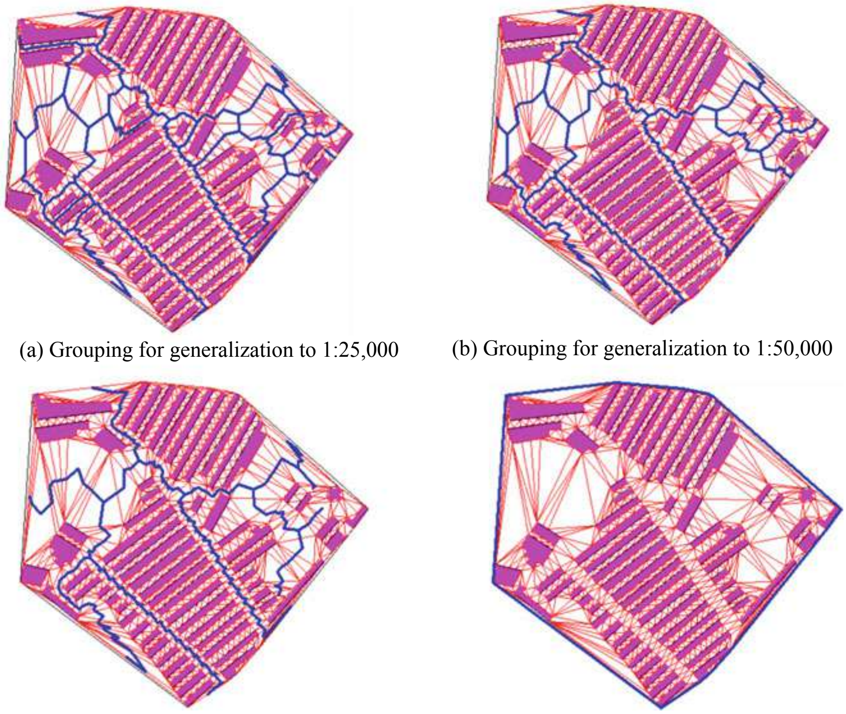
Section 8.3.5 described how a hierarchy of areas could be structured by a minimum spanning tree. In that example, the centroid of a polygon was used to represent the polygon. However, if the polygon is thin and/or irregular, then the edge length is not necessarily a good measure for closeness. Densification of points along the polygon edge will make the problem simpler. Figure 8.37 shows such an example. Figure 8.38 shows the transformation of buildings into suitable representations at different scales.

Li (1994) argued that the transformation in scale should be better performed in raster space (because a scale reduction causes a space reduction and the raster format takes care of space) and proposed the use of techniques in mathematical morphology for transformation in scale. Li et al. have developed a complete set of algorithms for such transformations based on mathematical morphology.

One such algorithm is the aggregation of areas into groups and transformation into representations at different scales (Su et al. 1997). The mathematical model for the aggregation is:

$$\mathrm{C} = (\mathrm{A} \oplus \mathrm{B}_1) \ominus \mathrm{B}_2 \tag{8.20}$$

where A is the representation (image) showing the original area features and B_1 and B_2 are the two structuring elements.



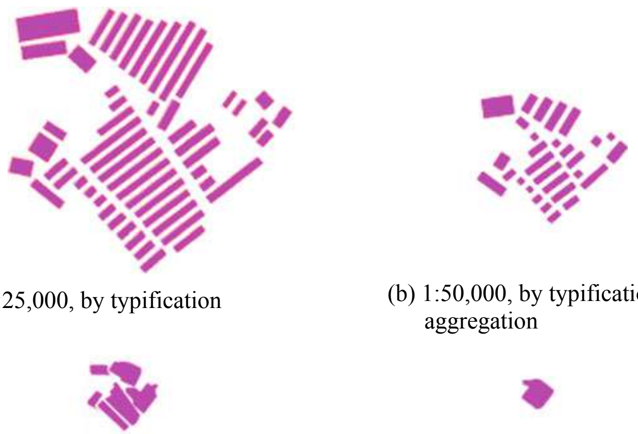
(a) Grouping for generalization to 1:25,000

(b) Grouping for generalization to 1:50,000

(c) Grouping for generalization to 1:100,000

(d) Grouping for generalization to 1:250,000

Fig. 8.37 Grouping of buildings at 1:10000 scale for generalization to various scales (Li et al. 2004)



(a) 1:25,000, by typification

(b) 1:50,000, by typification and aggregation

(c) 1:100,000, by aggregation

(d) 1:250,000, by aggregation

Fig. 8.38 Transformation of grouped buildings to various scales (Li et al. 2004)

The success of applying this model to area combination depends on the proper size and shape of the structuring elements B_1 and B_2. Su et al. (1997) suggest that the size of B_1 and B_2 should be determined by the input and output scales, following the Natural Principle described in Sect. 8.2.3. Figure 8.39 shows the combination of

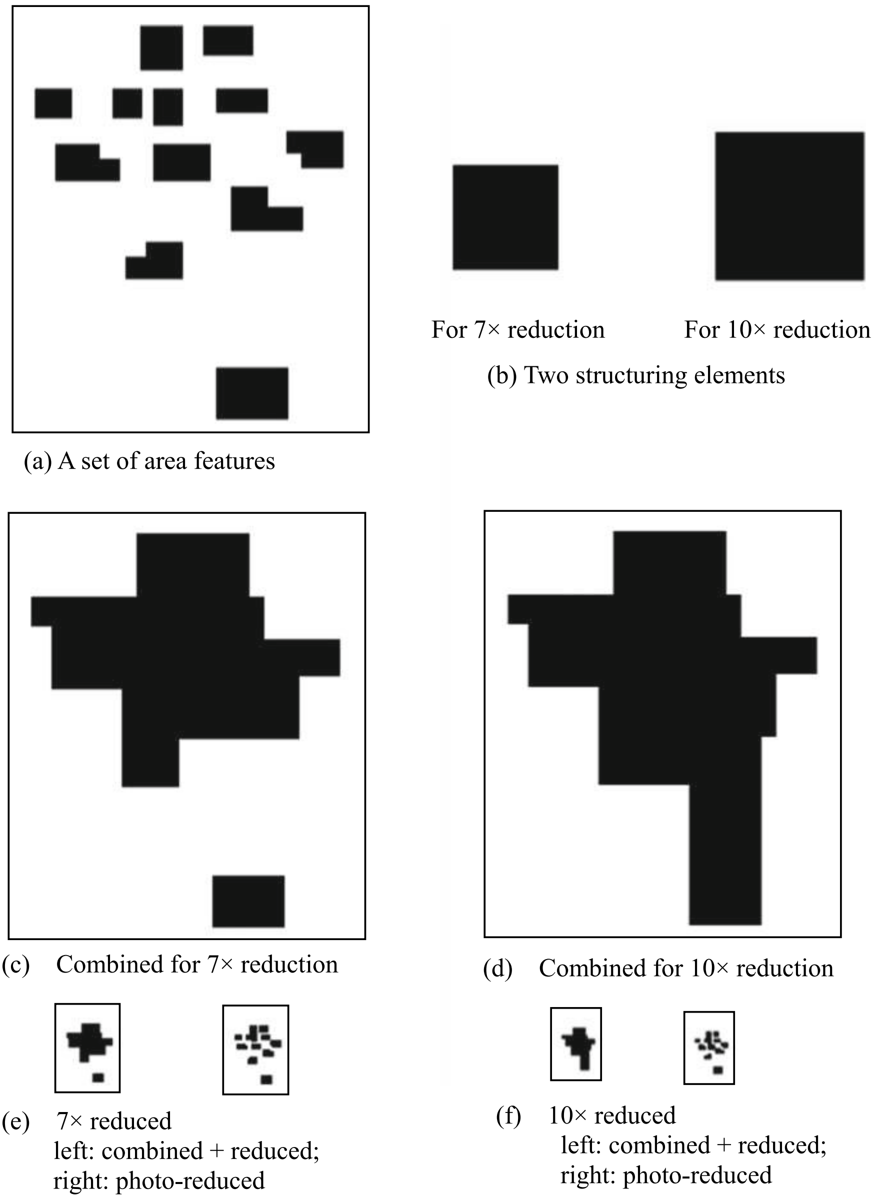


Fig. 8.39 Combination of area features at different scales (Extracted from Su et al. 1997)

(a) A settlement with an irregular shape (b) Simplified by the SLLM algorithm

Fig. 8.40 Shape refinement by the SLLM algorithm (Su et al. 1997)

buildings using this model for two different scales: one for a scale reduction by 7 times and the other by 10 times. The results are also compared with those using simple photoreduction. The combined results are very reasonable. However, the combined results are very irregular and the simplification of boundaries could be discussed. A detailed description of such a simplification is omitted here but can be found in the work of Su et al. (1997) and the book by Li (2007). The result is shown in Fig. 8.40.

8.4.7 Mathematical Solutions for Transformation (in Scale) of Spherical and 3D Features

In the previous sections, mathematical solutions for transformation of 2D features have been presented. Mathematical solutions for transformation of spherical (e.g., Dutton 1999) and 3D features (e.g., Anders 2005) have also been researched, although the body of literature is much smaller than that for map generalization. In recent years, there have been more papers on the generalization of buildings-based CityGML (e.g., Fan and Meng 2012, Uyar and Ulugtekin 2017); details on such methodologies are omitted here due to page limitations.

8.5 Transformation in Scale: Final Remarks

The beginning of this chapter emphasized that continuous zooming is at the core of Digital Earth as initiated by Al Gore. Continuous zooming is a kind of transformation of spatial representation in scale. In this chapter, the theoretical foundation for transformations in scale was presented in Sect. 8.2. Then, models for such transformations were described in Sect. 8.3 for raster and vector data, images, digital terrain models and map data. A selection of algorithms and/or mathematical functions for achieving these transformations was presented in Sect. 8.4.

Notably, the content of this chapter was concentrated on the theories and methodology to achieve continuous zooming and some important issues related to transformation in scale have been omitted, such as temporal scale, scale effect and optimum scale selection. For the content of the models for transformation in scale, emphasis was on the representations. Thus, other models such as geographical and environmental processes were excluded. However, these aspects are important but were omitted due to page limitations.

References

Abler R (1987) The National Science Foundation National Center for Geographic Information and Analysis. International Journal of Geographical Information Systems 1(4): 303–326.

Ai TH, Guo R Z (2007) Polygon Cluster Pattern Mining Based on Gestalt Principles. Acta Geodaetica et Cartographica Sinica 36(3): 302–308.

Anders KH (2005) Level of detail generation of 3D building groups by aggregation and typification. Paper presented at the XXII International Cartographic Conference, A Coruña, Spain.

Anderberg M R (1973) Clustering Analysis for applications. Academic Press, New York. p 376.

Atkinson, P.M., 2008. Super-resolution mapping using the two-point histogram and multi-source imagery. In: geoENV VI – Geostatistics for Environmental Applications (Part four: remote sensing). Springer, Netherlands, 307–321.

Atkinson PM (2013) Downscaling in remote sensing. International Journal of Applied Earth Observation and Geoinformation 22: 106–114.

Ball G, Hall D (1967) A clustering technique for summarizing multivariate data. Behavioral Science 12(2): 153–155.

Boucher A, Kyriakidis PC (2006) Super-resolution land cover mapping with indicator geostatistics. Remote sensing of environment 104(3): 264–282.

Burt P, Adelson E (1983) The Laplacian pyramid as a compact image code. IEEE Transactions on communications 31(4): 532–540.

Chaudhry O, Mackaness WA (2005) Rural and urban road network generalisation deriving 1:250, 000 from OS mastermap. In: Proceedings of the 22th International Cartographic Conference, La Coruña. July 10–16, 2005.

Chen J, Hu YG, Li ZL, Zhao RL, Meng LQ (2009) Selective omission of road features based on mesh density for automatic map generalization. International Journal of Geographical Information Science 23(8):1013–1032.

Clarke K, (1995) Analytical and Computer Cartography. Prentice Hall. 334 pp.

Douglas D, Peucker T (1973) Algorithms for the reduction of the number of points required to represent a digitized line or its caricature. The Canadian Cartographer 10(2): 112–122.

Dutton GH (1981) Fractal Enhancement of Cartographic Line Detail. American Cartographer 8(1): 23–40.

Dutton GH (1984) Part 4: Mathematical, Algorithmic and Data Structure Issues: Geodesic Modelling of Planetary Relief. Cartographica 21(2–3): 188–207.

Dutton GH (1996) Encoding and Handling Geospatial Data with Hierarchical Triangular Meshes, In: Kraak M J Molenaar M (Eds), In: Proceeding of 7th International Symposium on Spatial Data Handling. Netherlands, pp 34–43.

Dutton GH (1999). Scale, sinuosity and point selection in digital line generalization. Cartography and Geographic Information Science. 26(1): 33–53.

Estivill-Castro V, Lee I (2002) Multi-level clustering and its visualization for exploratory spatial analysis. GeoInformatica 6: 123–152.

Fan H, Meng L (2012) A three-step approach of simplifying 3D buildings modeled by CityGML. International Journal of Geographical Information Science 26(6):1–18.

Goovaerts P (2008) Kriging and semivariogram deconvolution in presence of irregular geographical units. Mathematical Geosciences 40(1): 101–128.

Goovaerts P (2009) Combining area-based and individual-level data in the geostatistical mapping of late-stage cancer incidence. Spatial and Spatio-temporal Epidemiology 1(1): 61–71.

Goovaerts P (2010) Combining areal and point data in geostatistical interpolation: applications to soil science and medical geography. Mathematical Geosciences 42(5): 535–554.
Gore A (1998) The Digital Earth: Understanding Our Planet in the 21st Century. Speech given at the California Science Center, Los Angeles, California, on January 31, 1998.
Jiang B, Claramunt C (2004) A Structural Approach to the Model Generalization of urban Street Network. GeoInformatica 8(2): 157–173.
Kolbe TH, Gröger G, Czerwinski A, Nagel C (2008) OpenGIS City Geography Markup Language (CityGML) Encoding Standard. Technical Report OGC 08-007rl, Open Geospatial Consortium Inc. http://www.opengeospatial.org/standards/citygml. Accessed 1 July 2019.
Kyriakidis P (2004) A Geostatistical framework for area-to-point spatial interpolation. Geographical Analysis 36(3): 259–289.
Lee YC, Li ZL, Li YL (2000) Taxonomy of space tessellation. ISPRS Journal of Photogrammetry and Remote Sensing 55(2000): 139–149.
Li ZL (1994) Mathematical morphology in digital generalization of raster map data. Cartography 23(1): 1–10.
Li ZL (2005) Digital Terrain Modeling: Principles and Methodology. Boca Raton: CRC Press. p 340.
Li ZL (2007) Algorithmic Foundations of Multi-Scale Spatial Representation. Boca Raton: CRC Press. p 281.
Li ZL (2008) Multi-Scale Digital Terrain Modelling and Analysis. In: Zhou Q, Lees B, Tang G. (eds), Advances in Digital Terrain Analysis, Springer, pp 59–83.
Li ZL, Openshaw S (1992) Algorithms for objective generalization of line features based on the natural principle. International Journal of Geographical Information Systems 6(5): 373–389.
Li ZL, Openshaw S (1993) A natural principle for objective generalisation of digital map data. Cartography and Geographic Information Systems 20(1): 19–29.
Li ZL, Li C (1999) Objective algorithms for multi-scale representation of relief. Proceedings of 2rd International Workshop on Dynamic and Multi-dimensional GIS. 4–6 Oct. 1999, Beijing. pp 17–22.
Li ZL, Zhou Q (2012) Integration of linear- and Areal-hierarchies for continuous multi-scale representation of road networks. International Journal of Geographical Information Science 26(5): 855–880.
Li ZL, Yan H, Ai T, Chen J (2004) Automated Building Generalization Based on Urban Morphology and Gestalt Theory. International Journal of Geographical Information Science 18(5): 513–534.
Li ZL, Liu Q, Tang J (2017) Towards a scale-driven theory for spatial clustering. Acta geodetica et cartographica sinica 46(10): 1534–1548.
Mackaness W, Ruas A, Sarjakoski LT (eds) (2007) Generalisation of geographic information: Cartographic modelling and applications. Elsevier Science. p 386.
Mallat S (1989) A theory for multiresolution signal decomposition: the wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence 11(7): 674–693.
McMaster RB, Shea KS (1992) Generalization in Digital Cartography. Association of American Geographers, p 134.
Miller H, Han J (2009) Geographic Data Mining and Knowledge Discovery (2nd edition.). CRC Press, New York. p 486.
OGC (Open Geospatial Consortium) (2019) Discrete Global Grid Systems DWG. http://www.opengeospatial.org/projects/groups/dggsdwg. Accessed 1 July 2019.
Openshaw S (1984) The Modifiable Areal Unit Problem. CATMOG # 38. Norwick, England: Geo Books.
Openshaw S, Charlton M, Wymer C, Craft A (1987) A mark 1 geographical analysis machine for the automated analysis of point data. International Journal of Geographical Information Systems 1: 335–358.
Quattrochi DA, Goodchild MF (eds) (1997) Scale in Remote Sensing and GIS. CRC Press, Boca Raton. p 406.

Raposo P (2013) Scale-specific automated line simplification by vertex clustering on a hexagonal tessellation. Cartography and Geographic Information Science 40(5): 427–443.
Rhind D (1988) A GIS research agenda. International Journal of Geographical Information Systems 2(1): 23–28.
Robinson AH, Sale R, Morrison JL, Muehrcke PC (1984) Elements of Cartography. (5th Edition), Wiley, New York.
Sheppard E, McMaster R (eds.) (2004) Scale and Geographic Inquiry: Nature, Society and Method. Blackwell Publishing, Malden. p 272.
Sposito G (ed.) (1998) Scale Dependence and Scale Invariance in Hydrology. Cambridge University Press. p 438.
Stefanakis E (2017) Web Mercator and raster tile maps: two cornerstones of online map service providers. Geomatica 71(2):100–109.
Stewart JB, Engman ET, Feddes RA, Kerr Y (1996). Scaling up in Hydrology using Remote Sensing, John Wiley & Sons, Chichester. p 255.
Su B, Li ZL, Lodwick G, Muller JC (1997) Algebraic models for the aggregation of area features based on morphological operators. International Journal of Geographical Information Systems 11(3): 233–246.
Tan ST (2018) Morphology-based Modelling of Aggregation Effect on the Patch Area Size and its Application. PhD Thesis. Southwest Jiaotong University, China.
Tate N, Atkinson P (2001) Modelling Scale in Geographical Information Science. John Wiley & Sons, Chichester. p 277.
UCGIS (2006) UCGIS Research Agenda. https://gistbok.ucgis.org/publication/research-priorities Accessed 10 July 2018.
Uyar A, Ulugtekin NN (2017) A proposal for generalization of 3D models. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume IV-4/W4. https://doi.org/10.5194/isprs-annals-IV-4-W4-389-2017. Accessed 10 July 2018.
Visvalingham M, Whyatt J (1993) Line generalization by repeated elimination of points. The Cartographic Journal 30(1): 46–51.
Wang Q, Shi WZ, Atkinson PM, Zhao Y (2015) Downscaling MODIS images with area-to-point regression kriging. Remote Sensing of Environment 166: 191–204.
Weibel R (1996) A typology of constraints to line simplification. In: Proceedings of 7th International Symposium on Spatial Data Handling (SDH'96), pp 9A.1–9A.14.
Weng Q (ed) (2014) Scale Issues in Remote Sensing. John Wiley & Sons, Hoboken. p 334.
Woodcock C E, Strahler A H T (1987) The factor of scale in remote sensing. Remote Sensing of Environment 21: 311–332.
Xu R, Wunsch IID (2005) Survey of clustering algorithms. IEEE Transactions on Neural Networks 16: 645–678.
Xu Y, Huang B (2014) A spatio-temporal pixel-swapping algorithm for sub-pixel land cover mapping. IEEE Geoscience and Remote Sensing Letters 11(2): 474–478.
Zhang H, Li ZL (2009) Structural hole analysis for structuring hierarchical road networks. In: Proceedings of ICC2009, CD_ROM, p 11.
Zhang H, Li ZL (2011) Weighted Ego Network for Forming Hierarchical Structure of Road Networks. International Journal of Geographical Information Science 25(2): 255–272.
Zhang J, Atkinson P, Goodchild MF (2017) Scale in Spatial Information and Analysis. CRC Press.
Zhou Q, Li ZL (2012) A comparative study of various strategies to concatenate road segments into strokes for map generalization. International Journal of Geographical Information Science 26(4): 691–715.
Zhu KP, Wu F, Zhu Q (2007) Improvement and Assessment of Li-Openshaw Algorithm. Acta Geodaetica et Cartographica Sinica 36(4): 450–456.

Zhilin Li is the Chair Professor of Geo-informatics at the Hong Kong Polytechnic. His research interests include multi-scale modelling and representation, cartographic language and cartographic information theory and mapping from satellite images. He was a vice president of the International Cartographic Association.

Haowen Yan is a professor in geographic information science at Lanzhou Jiaotong University, China. His research interests are map generalization, spatial relations and spatial data security. He is now one of the three editors-in-chief of Journal of Geovisualization and Spatial Analysis published by Springer-Nature.

Chapter 9
Big Data and Cloud Computing

Yun Li, Manzhu Yu, Mengchao Xu, Jingchao Yang, Dexuan Sha, Qian Liu and Chaowei Yang

Abstract Big data emerged as a new paradigm to provide unprecedented content and value for Digital Earth. Big Earth data are increasing tremendously with growing heterogeneity, posing grand challenges for the data management lifecycle of storage, processing, analytics, visualization, sharing, and applications. During the same time frame, cloud computing emerged to provide crucial computing support to address these challenges. This chapter introduces Digital Earth data sources, analytical methods, and architecture for data analysis and describes how cloud computing supports big data processing in the context of Digital Earth.

Keywords Geoscience · Spatial data infrastructure · Digital transformation · Big data architecture

9.1 Introduction

Digital Earth refers to the virtual representation of the Earth we live in. It represents the Earth in the digital world from data to model. Data are collected and models are abstracted to build the digital reality. Massive amounts of data are generated from various sensors deployed to observe our home planet while building Digital Earth. The term "big data" was first presented by NASA researchers to describe the massive amount of information that exceeds the capacities of main memory, local disk, and even remote disk (Friedman 2012). According to the National Institute of Standards and Technology (NIST), "*Big Data is a term used to describe the large amount of data in the networked, digitized, sensor-laden, information-driven world*" (Chang and Grady 2015). This definition refers to the bounty of digital data from various data sources in the context of Digital Earth, which focus on big data's geographical aspects of social information, Earth observation (EO), sensor observation service (SOS), cyber infrastructure (CI), social media and business information (Guo 2017; Guo et al. 2017; Yang et al. 2017a, b). Digital Earth data are collected from satellites,

Y. Li · M. Yu · M. Xu · J. Yang · D. Sha · Q. Liu · C. Yang (✉)
Department of Geography and GeoInformation Science, NSF Spatiotemporal Innovation Center, George Mason University, Fairfax, VA, USA
e-mail: cyang3@gmu.edu

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_9

Table 9.1 Definition of the "9Vs" of big data

"V"	Definition
Volume	The vast data size that traditional data storage and computing technologies cannot easily capture, store, manipulate, analyze, manage and present
Variety	The diversity of data formats and sources. The data formats include text, geometries, images, video, sounds or a combination
Velocity	The speed of data production, storage, analysis, and visualization based on advanced development of data collection methods, i.e., the massive number of sensors in the Interest of Things (IoT) and social media networks
Veracity	The varying reliability, accuracy, or quality of data sources
Validity	The accuracy and correctness of Earth data for the intended usage
Variability	The meaning of data continues to change, particularly for Earth data that relies on natural language processing
Vulnerability	Data security is an important part of typical and big Earth data because some geospatial data contain identification information related to people or governments
Volatility	The timeliness and freshness of Earth data
Visualization	Visualization of Earth data is challenging with limited memory, poor scalability and functionality, and various data increasing at a high velocity
Value	Value reflects the tremendous straightforward and potential scientific and social worth based on imaginative insight and analysis results

sensors, simulation models, mobile phones, utilities, vehicles, and social networks in different formats, e.g., imagery, text, video, sound, geometries and combinations of them (Yang et al. 2017a, b). Digital Earth data are naturally big data because of the variety of data sources and enormous data volume.

The increasing availability of big Earth data has provided unprecedented opportunities to understand the Earth in the Digital Earth context. In recent research, big data have been characterized by 5 Vs (volume, variety, velocity, veracity, and value) (Gantz and Reinsel 2011; Zikopoulos and Barbas 2012; Marr 2015). Firican (2017) extended the 5 Vs into big data characteristics including variability, validity, vulnerability, volatility and visualization (as defined in Table 9.1 and further elaborated below).

Volume
The volume of remote sensing imagery collected by satellites and drones easily reaches the TB and PB levels. For example, the Integrated Multi-satellite Retrievals for GPM (IMERG) data product records global precipitation information every half hour, producing up to 3.45 TB data yearly (Huffman et al. 2015). Other location-based data such as social media (e.g., Twitter) and VGI (e.g., OpenStreetMap) are constantly growing.

Variety
Data sources include sensors, digitizers, scanners, numerical models, mobile phones, the Internet, videos, emails, and social networks in the context of Digital Earth. All types of geospatial data require a more effective data structure, framework, index, model, management methodology, and tactics. In addition, these geospatial data are formatted in various data models, e.g., vector and raster, structured and unstructured.

Velocity
The speed of Earth data collection and generation has increased with the development of advanced techniques such as drone observation for disaster monitoring. For example, with the massive number of object-based sensors in the IoT, the data generation of IoT nodes is fast since most sensors continuously generate data in real time.

Veracity
The accuracy of geospatial data varies by data source (Li et al. 2016). Taking precipitation as an example, the quality of remote sensing images such as TRMM and IMERG depends on the sensor configuration, calibration methods, and retrieval algorithms. Precipitation information in MERRA (Modern Era Retrospective-analysis for Research and Applications) data relies on the sophistication of meteorological models. Stationary data collected by rain gauges are more accurate even though they are sparse.

Validity
Similar to veracity, validity concerns the accuracy and correctness of Earth data for the intended usage. In addition to data selection in which data are chosen with appropriate spatial and temporal resolutions and variables for a specific application, data preprocessing, e.g., data augmentation, interpolation, outlier detection, also play an important role in uncovering information from big Earth data. Consistent data quality, common definitions and metadata can benefit the community, resulting in Earth data of high validity.

Variability
Variability refers to the continuous change in the meaning of data in the context of big Earth data, particularly for data that relies on natural language processing. For example, Twitter data emerged as an additional source for natural disaster management (Yu et al. 2018), as tweets posted during disasters can be collected to aid situational awareness. The meaning of words constantly changes over time, for example, the word "Irma" may be a name but started to represent the strongest observed hurricane in the Atlantic in most tweets around October 2017.

Vulnerability
Security is a challenging aspect because some geospatial data contain identifiable information or are sensitive. For example, cellular data have been widely utilized to analyze human activities in smart city applications, however, showing phone numbers may divulge people's private affairs.

Volatility
Volatility refers to the timeliness and freshness of Earth data, i.e., how long the Earth data stay useful and relevant to applications and how long the data should be kept. Due to the velocity and volume of big Earth data, it is impossible to store all the data in a live database without any performance issues. A series of rules should be established for data currency, availability and rapid retrieval (Firican 2017), e.g., historical and less frequently visited Earth data could be archived on a lower-cost tier of storage.

Visualization
Visualization of Earth data is a challenging task with limited memory due to poorly scalable, low-functionality, and high-velocity datasets. Traditional methods may fail to render billions of points, polylines and polygons when visualizing geospatial vector data, therefore graphical methods, e.g., data clustering, parallel coordinates, cone tree or circular network diagrams, should be used to represent Earth data (Firican 2017).

Value
Value presents a low-density pattern in the current big data ecosystem where only a small portion of data is utilized in practice. Earth data occupies 80%+ of our data assets (Dempsey 2012), but most datasets are not excavated and are under-utilized. With appropriate spatiotemporal resolution and analysis methods, the 9Vs have been addressed to obtain actionable knowledge to increase the value of big data.

Data collection strategies, data storage facilities, data analysis methods, and data access services facilitate the transformation from the other 9Vs to the 10th V of value. With the continuing increases in the volume and complexity of data, there are challenges in the life cycle of data management, including data storage, data query, data analysis, data sharing, and many other aspects. Managing big data requires an extensible, interoperable and scalable architecture that supports data storage and analysis. Fortunately, recent years have witnessed the evolution of cloud computing, which brings potential solutions to support the life cycle of big data management.

Cloud computing is a new computing paradigm for delivering computation as a fifth utility, which became popular earlier than big data (Yang et al. 2011a). It has the features of elasticity, pooled resources, on-demand access, self-service and pay-as-you-go characteristics (Mell and Grance 2011) and was termed spatial cloud computing in the context of Digital Earth (Yang et al. 2011a). Big data technologies, e.g., big data storage and big data analytics, evolve and benefit significantly from their integration with cloud computing.

To provide a comprehensive overview of how cloud computing supports big data in the context of Digital Earth, this chapter introduces Digital Earth data sources (Sect. 9.2), data analysis methods (Sect. 9.3), architecture for big data analysis (Sect. 9.4), and cloud computing and its support of big data management (Sect. 9.5). Two examples of EarthCube and Data Cube are introduced in Sect. 9.6 to exemplify cloud-based big data frameworks in the Digital Earth context.

9.2 Big Data Sources

With the advanced developments in Earth observation systems, various Earth data have been gathered at a high velocity from five major sources: (1) remote sensing, (2) in situ sensing, (3) simulation, (4) social media, and (5) infrastructure management (Fig. 9.1). Each covers more than one characteristic of big data. This section discusses the five data sources.

Remote sensing data
Referring to the USGS's definition, remote sensing is the process of detecting and monitoring the physical characteristics of an area by measuring the reflected and

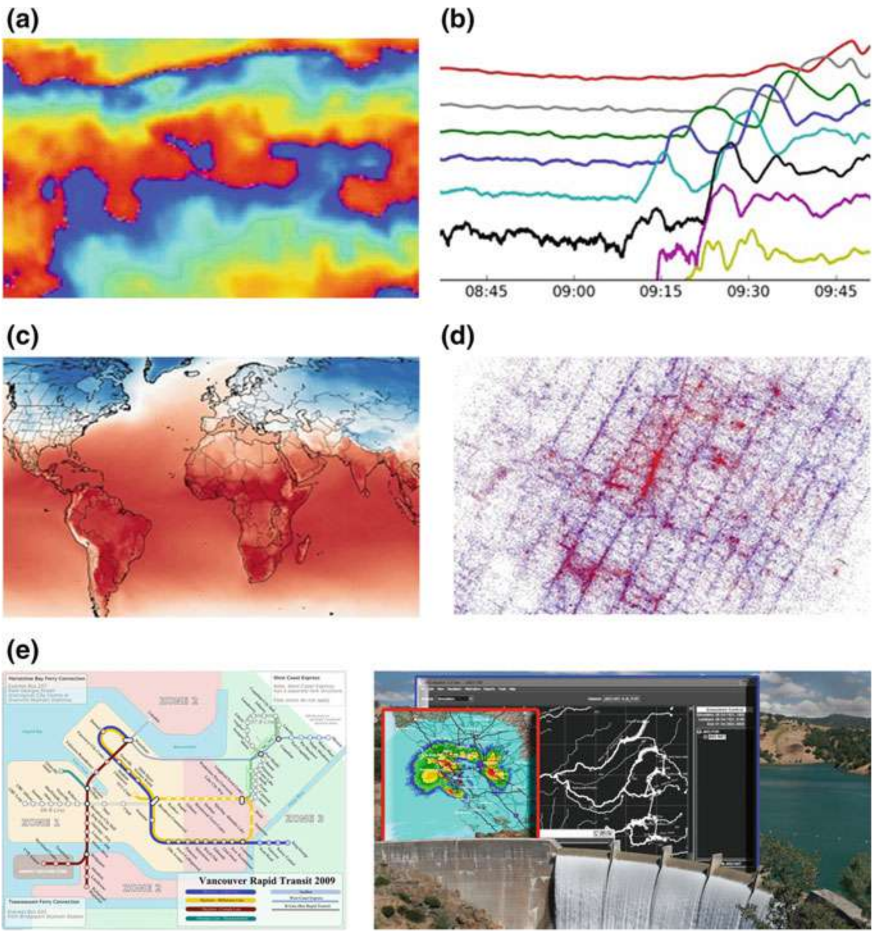


Fig. 9.1 Big Earth data sources: **a** remote sensing data (JPL 2001); **b** in situ data (NOAA 2017); **c** simulation data (Lipponen 2017); **d** social media data (Gundersen 2013); and **e** infrastructure data (Canada Line Vancouver Transit Map 2019; Robert 2000)

emitted radiation at a distance from the targeted area (USGS 2019). Such remotely observed data serve as a vital source for tracking natural phenomena, the growth of a city, changes in farmland or forest, and discovery of the rugged topography of the ocean floor. According to the Earth Observing System Data and Information System (EOSDIS) 2014 statistics, EOSDIS manages over 9 PB of data and adds 6.4 TB of data to its archives every day (NASA 2016). As data precision and density increase over time, data volume increases exponentially. In addition, ongoing international initiatives monitor the Earth in near-real time using satellites to support rapid data collection for quick and effective emergency response (Zhang and Kerle 2008). Remote sensing data are big data due to the big volume, variety, veracity and volatility.

In situ data
According to NOAA (National Oceanic and Atmospheric Administration), in situ data are measurements made at the actual location of the object. In contrast to remote sensing, in situ sensing harvests data directly at the observation location, and often provides continuous data streams to reflect the actual situation with very low latency. Examples of such measurements are (1) tall tower networks (NOAA ESRL/GMD) that provide regionally representative measurements of carbon dioxide (CO_2) and related gases and (2) moored and drifting buoys for marine/ocean data collection. In situ data are big data considering the volume, velocity, and veracity.

Simulation data
Simulation datasets or reanalysis datasets refer to the outputs of Earth models (e.g., climate) based on geophysical principles. By assimilating observations with models, better initial conditions can be leveraged and simulation results can be significantly improved, especially for short-term predictions. Simulation datasets can be used in various applications. For example, the precipitation, evaporation, and runoff from MERRA datasets can drive river flow models and enhance the study of sensitive ecosystems such as estuaries (Rienecker et al. 2011). In addition, the reanalysis winds used in transport models support the evaluation of aerosols. Simulation data are big data due to its volume, variety and validity.

Social media data
In recent years, social media has become one of the most popular sources of big data and provides valuable insights on event trends and people's references. Social networks such as Twitter and Facebook generate a vast amount of geo-tagged data every second and are transforming social sciences (Yang et al. 2017a). Scientists from economics, political science, social science, and geoscience domains utilize big data mining methods to detect social interactions and analyze health records, phone logs, and government records (Balakrishna 2012). For example, in Digital Earth, social media and crowdsourcing data can provide trends of the urban flooding events or wildfire spread, as well as support near-real time situational awareness when other types of data are limited or hard to obtain. However, social media data have high uncertainty and vary in format and quality. Tweet content analysis highly relies on natural language processing, but word meaning constantly changes. Social media

data are big data due to its volume, velocity, variety, veracity, validity, variability and volatility.

Infrastructure data
Infrastructure data serve as a vital data source of Digital Earth information, especially for developing smart cities. For example, basic infrastructure data (e.g., utility, transportation, and energy), healthcare data and governance data (e.g., environmental and construction management) should be proposed, planned and provided by local official departments and business agencies for a smart city (Hashem et al. 2016). Some infrastructure data may contain sensitive information. Taking water distribution management systems as an example, a synthetic data methodology was proposed to reproduce water consumption data according to privacy constraints (Kofinas et al. 2018). With the upgrades in infrastructure, Internet of Things (IoT) data, geo-tagged or geo-referenced data are continuously produced by various devices, sensors, systems and services (Boulos and Al-Shorbaji 2014). In the near future, various applications based on IoT data will benefit individuals and society. For example, near-real time data including temperature and wind information gathered by IoT sensors could support real-time urban microclimate analysis (Rathore et al. 2017). Infrastructure data are big data due to its volume, velocity, variety, veracity, vulnerability, validity and volatility.

Earth data are continuing to grow in volume and complexity. Big data analytical methods are utilized to mine actionable knowledge from big Earth data to convert the 9Vs of Earth data to the 10th V, which is discussed in the next section.

9.3 Big Data Analysis Methods

The advancements in remote sensing, social networking, high-performance simulation modeling and in situ monitoring provide unprecedented big data about our planet. The large volume and variety of data offer an opportunity to better understand the Earth by extracting pieces of knowledge from these data. This section discusses data analysis methods from the three aspects of data preprocessing, statistical analysis and nonstatistical analysis. The characteristics, applications, and challenges of these methods are introduced below.

9.3.1 Data Preprocessing

Real-world data are usually incomplete, noisy and inconsistent due to data collection limitations and sensor issues. Raw data may contain errors or outliers, lack specific attributes or have discrepancies in the descriptions. Therefore, data preprocessing (e.g., data cleaning, fusion, transformation, and reduction) are required to remove noise, correct data, or reduce data size.

Low-quality values (missing values, outliers, noises, inconsistent values) in raw data are often removed or replaced with user-generated values, e.g., interpolation values. The missing value is usually represented with a symbol (e.g., N/A) in raw data and easily recognize. Outliers and inconsistent values are hidden in the raw data and can be detected through statistical analysis. Taking water usage behavior analysis as an example, data preprocessing is necessary to turn smart water meter data into useful water consumption patterns because IoT sensors may fail to record data (Söderberg and Dahlström 2017).

Data transformation also plays an essential role in data preprocessing. Multiple Digital Earth data, e.g., climate data, soil moisture data, crop data, are converted to monthly z-score data before analysis to eliminate the seasonal trends that usually make the patterns of interest undiscoverable. Aggregation, another important data transformation method, groups data based on numerical attributes (Heuvelink and Pebesma 1999). In the Earth science domain, aggregating raw data to the county or state levels could uncover essential patterns for decision making, urban planning, and regional development.

Another trend in Digital Earth data analysis is multisource data fusion, which provides comprehensive data retrieved from several data sources. Generally, vector and raster data store Earth information with different spatial-temporal resolutions; thus, data must be converted to the same resolution by interpolating the lower resolution data or aggregating the higher resolution data for intelligent analysis to investigate scientific questions at a specific scale. Sharifzadeh and Shahabi (2004) introduced a spatial aggregation method that takes the sensor data distribution into account. Spatial interpolation is interpolation of point and areal data. Point interpolation is applied to contour mapping and areal interpolation is used in isopleth mapping (Lam 1983). In addition to spatial interpolation, temporal interpolation predicts values between timestamps (Lepot et al. 2017).

9.3.2 Statistical Analysis

In the era of big data, statistical analysis is a common mathematical method of information extraction and discovery. Statistical methods are mathematical formulas, models, and techniques used to find patterns and rules from raw data (Schabenberger and Gotway 2017). Data mining is the process of discovering patterns from large datasets involving statistical analysis. Through data mining, historical data can be transformed into knowledge to predict relevant phenomena. Both traditional statistics and data mining methods are discussed in this section. These methods include but are not limited to regression analysis, spatiotemporal analysis, association rules, classification, clustering, and deep learning.

Regression analysis
Regression models the relationships between a dependent variable and one or more explanatory variables (Yoo et al. 2014; Anderson 2015) by estimating the values of

a dependent variable when the values of the independent variables are known and the relationships exist. Regression models describe the strength or weakness of the relationship between several variables. For example, Blachowski (2016) proposed a weighted spatial regression method that identified four significant factors inducing land subsidence: thickness, inclination, the depth of coal panels, and the slope of the surface. There are challenges in spatial data analysis using regression methods, especially for situations that are complicated enough to result in serious residuals in the regression models.

Spatiotemporal analysis
Spatiotemporal data analysis investigates the trajectories and trends of spatiotemporal data. Li et al. (2016) investigated the spatiotemporal trends in the fluctuations of housing price data. Spatial data analytics and modeling techniques were used to identify the spatial distribution of housing prices at the micro level and explore the space-time dynamics of residential properties in the market, as well as the detected geographic disparity in terms of housing prices. Rahman and Lateh (2017) analyzed the temperature and rainfall time series data from 34 meteorological stations distributed throughout Bangladesh over 40 years (1971–2010) to statistically evaluate the magnitude of temperature and rainfall changes across space and time. Spatiotemporal analysis is still in its initial stage of development. Challenging questions remain, such as what kinds of patterns can be extracted from time series data and which methods and algorithms should be applied.

Association rule
Association rule learning is the process of discovering strong relationships between variables, i.e., rules, in a large database using measurements of support and confidence (Agrawal et al. 1993). In Digital Earth, Yang (2011b, 2016) applied association rules to mine the variables of Atlantic hurricanes from 1980 to 2003 and discovered a combination of factors related to rapid intensification probability, the low vertical shear of the horizontal wind (SHRD = L), high humidity in the 850–700 hPa range (RHLO = H), and tropical cyclones in an intensification phase (PD12 = H). Compared with traditional statistical methods, the rule-based mining method can find combinations of factors instead of a single factor related to an event.

Classification
Classification learning is the task of mapping input variables to discrete output variables called labels or categories, for example, 'building' or 'road.' It is the process of recognizing, differentiating and understanding objects. Support Vector Machine (SVM) is a classical classification algorithm in which a kernel-based metric is used to differentiate objects. Jiang et al. (2018b) integrated the ranking support vector machine (RankSVM) model from the computer science community with ocean data attributes to support data ranking in ocean portals. An SVM model is also used to predict geological lithofacies from wireline logs.

Clustering
Clustering is the process of splitting a set of objects into closely related groups, and each group is regarded as a cluster. Objects falling in the same cluster are more

similar to each other than those in other clusters. In Digital Earth, clustering plays an important role in pattern analysis. Hong and O'Sullivan (2012) clustered empirical datasets in Auckland, New Zealand for ethnic residential cluster detection, which is useful to understand contemporary immigrants and ethnic minorities in urban areas. Zhao et al. (2017) proposed a method to generate urban road intersection models from low-frequency GPS trajectory data. These patterns identified from empirical data are crucial for urban transportation planning and management.

Deep learning
As a new paradigm of machine learning, deep learning has achieved remarkable success in discovery of implicit knowledge (LeCun et al. 2015). In Digital Earth, deep learning algorithms have been adopted to solve domain problems. For example, Guo and Feng (2018) used multiscale and hierarchical deep convolutional features to assign meaningful semantic labels to the points in a three-dimensional (3D) point cloud, which is essential for generating 3D models. Li and Hsu (2018) proposed a deep learning approach to automatically identify terrain features (i.e., sand dunes, craters) from remote sensing imagery. Compared with traditional induction-based approaches, the deep learning approach could detect diverse and complex terrain features more accurately and process massive available geospatial data more efficiently.

9.3.3 Nonstatistical Analysis

In addition to statistical analysis, nonstatistical analysis methods also play an essential role in helping us descriptively understand Earth phenomena. This section introduces two representative models in Digital Earth, linked data and 3D city modeling.

Linked data
Linked data are structured data in which datasets are interlinked in the collection, which is useful for semantic queries and reasoning (Bizer et al. 2011). With linked data, data are sharable, and the relationships among the data are recorded. Standard web technologies such as RDF (Resource Description Framework) provide a way to build shareable linked data. In Digital Earth, heterogeneous Earth data (multidisciplinary, multitemporal, multiresolution, and multilingual) can be integrated based on linked data principles for decision making and knowledge discovery (Vilches-Blázquez et al. 2014). For example, Mc Cutchan (2017) proposed a structure of embedding geographic data into linked data and forecasted spatial phenomena with associated rules extracted from the linked data.

3D city modeling
A trend in Digital Earth analysis is to build a real 3D model with the aid of a computer, especially for cities where most human activities occur. 3D models provide real three-dimensional information for analysis, going beyond simple visualization of 3D objects. 3D informatics has become a cornerstone for a series of city-related

applications such as urban planning, skyline analysis, crisis and disaster management, route selection and navigation (El-Mekawy 2010). For example, Amirebrahimi et al. (2016) assessed and visualized flood damage using 3D urban modeling and a building information model (BIM), improving the resilience of the community to floods using detailed 3D information.

Knowledge distillation from Earth data has demonstrated excellent improvements in our understanding of the planet we live. As Earth data increase faster than ever, state-of-the-art analysis methods should be developed to handle the increasingly complicated spatiotemporal data. In addition, an extensible, interoperable and scalable architecture is a prerequisite for massive geographic data analysis, and we present a big data analysis architecture in the next section.

9.4 Architecture for Big Data Analysis

To support Earth data access/query/analysis in a reasonable response time, it is crucial to build a sophisticated analytical platform with robust architecture to reveal insights from the data (Yang et al. 2017a, b). Generally, the architecture of analytical platforms consists of a data storage layer, a data query layer, a data processing layer, and a visualization layer (Fig. 9.2).

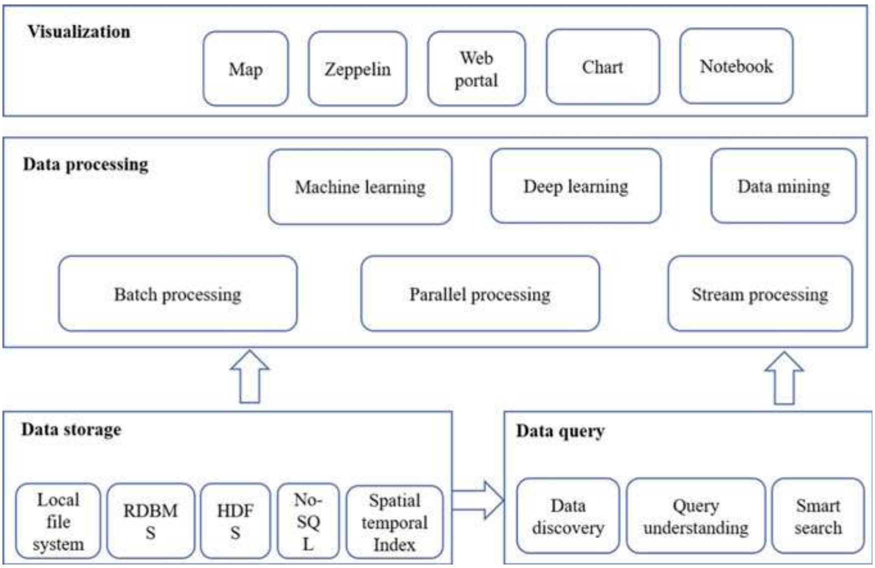


Fig. 9.2 Architecture for big data analyses

9.4.1 Data Storage Layer

Digital Earth is heavily involved in processing big data from Earth observations and model simulations, and these vast amounts of high-dimensional array-based spatiotemporal data pose challenges for data storage (Li et al. 2017b). Customizations are indispensable in integrating advanced data storage technologies with big Earth data storage, as general distributed file systems such as Hadoop are not designed to store spatiotemporal data.

A robust and stable data storage framework is the foundation of the data analysis architecture. A series of research efforts focused on optimizing spatiotemporal data storage in distributed file systems or databases. For example, Hu et al. (2018a) reorganized NetCDF (Rew and Davis 1990), a standard data format for array-based raster data, into CSV files and deployed them within SciDB (Cudre-Mauroux et al. 2009), a scalable multidimensional array clustering database. Zhao et al. (2010) converted NetCDF data into CDL (network Common data form Description Language) files and distributed them on HDFS (Hadoop Distributed File System). MERRA data, which store reanalysis Earth climatic variables in NetCDF format, were transformed into Hadoop Sequence Files to be processed by standard MapReduce functions (Duffy et al. 2012). Li et al. (2015) decomposed array-based raster data and stored them with HBase, a NoSQL database built upon HDFS in a cloud computing environment for efficient data access and query.

To enable efficient big data query, logical query capabilities have been proposed to support spatiotemporal query of array-based models such as SciHadoop (Buck et al. 2011). A spatiotemporal index was designed to efficiently retrieve and process big array-based raster data using MapReduce and a grid partition algorithm atop the index to optimize the MapReduce performance (Li et al. 2017a). SciHive was developed as an extension of Hadoop Hive, mapping arrays in NetCDF files to a table and calculating the value range for each HDFS to build a distributed adaptive index (Geng et al. 2013, 2014). Malik (2014) introduced a Z-region index into GeoBase to facilitate array-based data storage.

9.4.2 Data Query Layer

To help data consumers efficiently discover data from the massive available Earth data, the Digital Earth communities have built various data portals to improve the discovery, access, and usability of Earth data. The portals are normally supported by text search and spatiotemporal search and include the GeoPortal,[1] GeoNetwork[2] Spatial Web Portal (Xu et al. 2011), Global Earth Observation System of Systems GEOSS Clearinghouse (Liu et al. 2011; Nativi et al. 2015; Giuliani et al. 2017),

[1] https://www.geoportal.gov.pl/.

[2] https://geonetwork-opensource.org/.

GeoSearch (Lui et al. 2013) and many others. For example, GEOSS is a cloud-based framework for global and multidisciplinary Earth observation data sharing, discovery, and access (Nativi et al. 2015). In the framework, datasets or workflows are registered into shared collections or global catalogs, allowing for end users to search for workflows and datasets across multiple granularity levels and disciplines.

In addition, open-source information retrieval frameworks, e.g., Apache Lucene or its variants such as Solr and Elasticsearch (McCandless et al. 2010), were adopted to establish an Earth data portal instead of implementing a search engine for Earth data from scratch. Lucene uses the Boolean model and the practical scoring function to match documents to a query. Solr and Elasticsearch improve the Lucene index to enable big data search capabilities. The Physical Oceanography Distributed Active Archive Center (PO. DAAC) serves the oceanographic community with 514 collection level datasets and massive granule level data atop Solr (Jiang et al. 2018a). Elasticsearch is the fundamental component of NOAA's OneStop portal in which data providers manage data and metadata with increased discoverability and accessibility.

However, solely relying on open source solutions is insufficient for Earth data discovery because these solutions only rely on a keyword-based relevance score for ranking and ignore other user preferences, e.g., data processing level, sensor type. A few related research efforts have been conducted in the Earth science domain to make data search engines smarter and more intelligent. For example, an algorithm combing Latent Semantic Analysis (LSA) and a two-tier ranking was reported to build a semantic-enabled data search engine (Li et al. 2014a, b). Jiang et al. (2018a) developed a smart web-based data discovery engine that mines and utilizes data relevancy from metadata and user behavior data. The engine enables machine-learned ranking based on several features that can reflect users' search preferences.

9.4.3 Data Processing Layer

Data processing layer is a core component of the data analytics architecture. To analyze terabyte and petabyte datasets with low time latency, even in a real-time manner, sophisticated parallel computing algorithms and scalable computing resources are required in the big data processing framework (Yang et al. 2015a). Advanced open-source parallel computing solutions, e.g., Hadoop MapReduce, Spark, and their variants in the Earth data domain have been leveraged to support data analysis and mining tasks with better performance.

Hadoop MapReduce is a high-performance batch processing parallel framework that solves large computational problems on distributed storage systems (White 2012). It transfers the algorithm code to data nodes rather than moving data blocks to a compute node to avoid I/O bottlenecks. Spark enables high-performance data analysis with in-memory computing. An in-memory data structure called the Resilient Distributed Dataset (RDD) manages datasets distributed in a Spark cluster (Zaharia et al. 2012). However, the original distributed frameworks have limitations on big spatiotemporal data processing. Yu et al. (2015) noted that the system scalability for spatiotemporal data and interactive performance are the two main challenges for big Earth data processing. To solve these problems, scientists and engineers have customized open-source solutions for spatiotemporal data analysis.

SpatialHadoop is a MapReduce-based framework with native support for spatial data including a simple spatial high-level language, a two-level spatial index structure, a fundamental spatial component built on the MapReduce layer and three basic spatial operations (range query, k-NN query, and spatial link) (Eldawy and Mokbel 2015). GeoSpark provides operational tools for spatial big data processing based on Spark (Yu et al. 2015). The Spatial Resilient Distributed Datasets (SRDDs) structure represents spatial data blocks in memory and index objects using quad-tree and r-tree in each RDD partition (Lenka et al. 2016). ClimateSpark integrates Spark SQL and Apache Zeppelin to develop a web portal that facilitates interaction among climatologists, climate data, analytical operations and computing resources (Hu et al. 2018b). As an extension of Scala, GeoTrellis supports high-performance raster data analysis. GeoMesa provides spatiotemporal data persistence on Hadoop and column-family databases (e.g., Accumulo, HBase), as well as a suite of geospatial analytical tools for massive vector and raster data (Hughes et al. 2015).

As described in this section, a service-oriented, scalable architecture usually contains three major layers to provide desirable functionalities and capabilities: (1) the bottom data storage layer provides physical data storage, (2) the data query layer enables data discovery capabilities with proper functionality and interoperability, and (3) the data processing layer supports extensible, interoperable and scalable analytical functionalities based on open source solutions and their variants from the geoscience communities. With the architecture, big Earth data could be accessed and analyzed with low time latency or even in real time. However, it is challenging to set up such architecture and share data stored inside them due to the requirements of storage resources, computing resources, complicated configurations, and domain knowledge. Fortunately, the paradigm of cloud computing, discussed in the next section, brings potential solutions to ease the process of analytical framework setup and data sharing.

9.5 Cloud Computing for Big Data

9.5.1 Cloud Computing and Other Related Computing Paradigms

Grid computing and High Performance Computing (HPC) have been utilized for big data analytics. Grid computing, a distributed system of computer resources, performs large tasks using loosely coupled computers in a distributed system (Hamscher et al. 2000). The European Data Grid project utilizes grid computing to support exploration of multi-petabyte datasets (Segal et al. 2000) and the TeraGrid GIScience gateway utilized grid computing to perform computationally intensive geographical analytics

(Wang and Liu 2009). HPC uses supercomputers to run applications in parallel, efficiently and quickly, and is used in the PRACE project to serve European scientists with high-performance computing capabilities to conduct research (Hacker et al. 2010).

Cloud computing virtualizes computer resources as a resource pool to provide computing resources over the network by optimizing resource usage in terms of the CPU, RAM, network, and storage. Cloud computing has intrinsic connection to the Grid Computing paradigm (Foster et al. 2008) in that both are distributed computing systems. Cloud computing relies on remote servers whereas grid computing connects servers or personal computers over a common network using a Wide Area Network (WAN) to perform parallel tasks (Foster et al. 2008). Compared with HPC, cloud computing is cost effective and easy to use. Although cloud computing can provide high performance computing capability, HPC is irreplaceable for some applications since supercomputers are required to process very complicated processes such as climate simulations. In addition, resources in cloud computing are controlled by the service providers and users have limited controls.

In addition to cloud computing, other new computing paradigms have emerged to build a comprehensive and economic computing framework. For example, edge computing can process data at the edge of network due to the advancement of the IoTs. IoT applications usually produce a massive amount of streaming data and require near-real time response; thus, cloud computing alone is not an optimal solution for data collection and analysis for such real-time applications. In edge computing, edge nodes serve as data providers and consumers to protect data privacy and make full use of the computing capacity of edge nodes. Less data is transferred to the cloud computing platform after data preprocessing in edge nodes, reducing the response time and bandwidth cost (Shi et al. 2016).

Mobile computing with portable computing nodes has become an important computing paradigm with the improvements in the computing and storage capacity of smart devices such as smartphones and tablets (Qi and Gani 2012). Although services provided by mobile computing are not as reliable as edge computing and cloud computing due to the restrictions in battery volume and network connection, mobile computing can collect data and reach end users where cloud computing and edge computing are inaccessible.

These computing paradigms have advantages and disadvantages, and can be integrated to complement each other and provide reliable and effective data storage and processing frameworks according to the data characteristics and computing requirements. Cloud computing and big data are the two most important technologies in Digital Earth. The following section discusses the utilization of cloud computing to support big data management in Digital Earth.

9.5.2 *Introduction to Cloud Computing*

As a new computing paradigm, cloud computing delivers scalable, on-demand, pay-as-you-go access to a pool of computing resources (Mell and Grance 2011; Yang et al. 2011a). Practically, cloud computing aims to maximize the utilization rate of physical resources and provide virtual resources to aid applications and services. Cloud computing relies on several technologies including virtualization, network security, and high availability to provide services over the network. These technologies make it easier, more efficient, and more economical to set up architecture for big data analysis.

Virtualization is the fundamental technology for cloud computing, which abstracts an application, operating system, or data store from the underlying hardware or software. Virtualization creates a "layer of abstraction" between the physical systems and a virtual environment in the virtualization process (Big Data Virtualization). Server virtualization optimizes the use of redundant computing and storage resources by virtualizing distributed computer resources (e.g., CPU, RAM, Network, and Disk) and managing them in the same resource pool. With virtualization, cloud computing can provide on-demand big data services and support big data technologies including big data storage, process, analysis, visualization, and remote collaboration (Fig. 9.3). Virtualizing big data resources as a pool serves as a user-friendly interface and makes big data analytics accessible to end users.

As one of the cloud solutions, public clouds are the most accessible cloud computing services offered by third-party providers (e.g., Amazon Web Services (AWS), Microsoft Azure, Alibaba Cloud) over the Internet. Public clouds are available to the public and may be offered on a pay-per-usage model (Li et al. 2010). In contrast to public clouds, private clouds are dedicated for use inside an organization. Private cloud resources can be managed by an organization or by a third-party vendor,

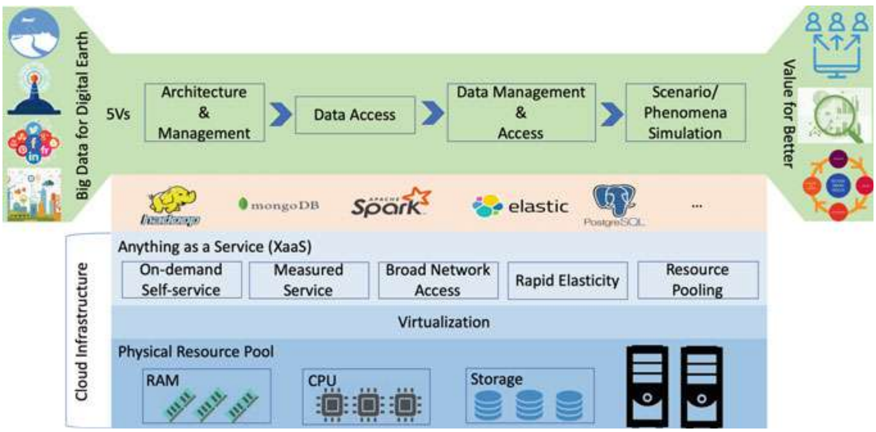


Fig. 9.3 Cloud computing for big data analysis

regardless of the physical location of the resources (Dillon et al. 2010). The computing resources in a private cloud are isolated and delivered via a secure private network.

The advanced features of auto scaling and load balancing through resource monitoring further maximize the capability of cloud computing resources. Based on the individual performance of a machine, autoscaling can be applied to allow for some servers to rest during times of low load to save electricity costs and automatically add more instances during times of high demand. In addition, load balancing improves the distribution of workloads across multiple instances to optimize resource use, minimize response time, and avoid overloading any single instance.

9.5.3 Cloud Computing to Support Big Data

Cloud computing combines distributed computing resources into one virtual environment, providing big data analytics and solutions during the life cycles of big data. Three main categories of cloud computing services are (1) Infrastructure as a Service (IaaS), (2) Software as a Service (SaaS), and (3) Platform as a Service (PaaS). Together with Data as a Service (DaaS), Model as a Service (MaaS; Li et al. 2014a, b) and workflow as a service (WaaS; Krämer and Senner 2015), cloud computing offers big data researchers the opportunity of anything as a service (XaaS; Yang et al. 2017b).

Cloud Storage for Big Data Storage

The characteristics of big data in high volume lead to challenges for data storage. Cloud computing's potential for unlimited storage support helps solve the volume challenge of big data, as the cloud provides virtually customizable storage with elastically expandable and reducible size. An alternative solution is Data Storage as a Service (DSaaS) enabled by block storage, which is the capability of adding external storages as "blocks". With block storage, it is possible to enlarge the storage size without physically loading hard drives. Virtually unlimited scalable storage offered by cloud computing grants users the capability of dynamic adjustment to satisfy the storage requirements of data with high volume and velocity. The modularized virtual resource offers effortless data sharing within production environments by allowing for an external data block to be detached and remounted from one machine to another. External data storage can be automatically backed up to prevent users from losing data, and backups that are securely saved at the back-end server can be easily transferred and restored. In addition, information security is guaranteed because the physical location cannot be obtained from the disk drive (Mayama et al. 2011).

Cloud Computing for Big Data Processing

Processing large volumes of data requires dedicated computing resources, e.g., faster CPUs and networks and larger disks and RAMs (Yang et al. 2017b). Cloud computing provides on-demand resources and delivers configurable resources including mountable external storage spaces, computing resources (CPU, RAM), and network services. Traditionally, a computer uses approximately two-thirds of the power of a busy computer (JoSEP et al. 2010), and cloud computing has the potential to provide on-demand computing resources. Isolated virtual structures have been created for big data systems to enhance system stabilities, which can be easily managed in different file systems and replicated through backup images to provide fast configuration recovery. The ability to replicate environments automates the expansion of compute nodes in virtual machine clusters, thereby efficiently utilizing resource pools to support big data analytics. With the foundational support of storage for big data, data processing inherits the advantages of fast data acquisition and relocation.

Although cloud computing could serve as an excellent infrastructure option for big data processing, several aspects should be considered to minimize the bottleneck effect for the general processing speed, such as the choice of cloud volume type according to I/O demand and cloud bandwidth selection according to application requirements.

Cloud Computing for Big Data Analytics

Popular big data analytical platforms such as Apache Hadoop are traditionally installed on physical machine clusters, resulting in a waste of computing resources due to hardware redundancy (CPU and RAM). With the virtual clusters provided by cloud computing through virtualization technology, distributed analytical platforms can be migrated to the virtual clusters from physical machine clusters, optimizing the usage of computing resources in an efficient manner.

With the aid of autoscaling and load balancing, deploying on-demand and scalable big data analytical platforms could easily provide resilient analytical frameworks and minimize waste of computing resources. Autoscaling supports parallel algorithms on distributed systems and architectures for scalability. It allows for the expanded resources to function when the algorithms or programs are enabled with parallel computing capability. Without it, public cloud providers such as AWS could not offer automatic scalability (JoSEP et al. 2010). The load balancer distributes workloads among virtual clusters and triggers autoscaling functions when analytics require higher computing configurations. The virtual system as a whole could dynamically fit higher computing requirements by launching more virtual duplications as needed. The load balancer acts as a virtual network traffic distributor and can be optimized to better allocate overall resources.

Cloud Computing for Big Data Sharing and Remote Collaboration

Traditional deployment of big data systems requires complicated settings and efforts to share data assets. It lacks access control and often leads to data security and data privacy issues. Cloud computing enhances the sharing of information by applying modern analytical tools and managing controlled access and security (Radke and Tseng 2015). Virtualization enables different parties to share data assets to achieve various goals and objectives under a centralized management system. With the support of cloud computing, it is possible to flexibly share data and remotely collaborate, which involve interdisciplinary collaborations and advanced workflows. Though data sharing, computational resource sharing, and production environment sharing, cloud computing can potentially be used to build a perceptual environment to support various businesses and applications (Li et al. 2015). Unfortunately, workflow sharing remains challenging due to domain boundaries (Yang et al. 2017b).

9.6 Case Study: EarthCube/DataCube

Big data and cloud computing enable Earth scientists and application developers to create web-accessible frameworks and platforms to efficiently store, retrieve and analyze big Earth data. In the Earth science domain, scientists have proposed a series of data models, frameworks, and initiatives to ensure the success of heterogeneous data sharing and analysis. For example, a 10-year framework initiative on sustainable consumption and production from 2013 to 2023 was launched by the United Nations Environmental Program (UNEP). The Future Earth framework, an international research program in the environmental science community, serves as an evolving platform to support transitions toward sustainability (Lahsen 2016). Microsoft's Eye-On-Earth platform aids climate change research in several European countries by collecting and sharing water and air quality data (Microsoft 2011). As part of the European program to monitor the Earth, the Copernicus Data and Information Access Service (DIAS) platform collects and processes data from remote and in situ sensors and provides reliable information covering six thematic areas including land, ocean, atmosphere, climate, emergency, and security (Bereta et al. 2019). Through its Exploitation Platforms (EP) initiative, the European Space Agency (ESA) built several cloud-based Thematic Exploitation Platforms (TEPs) in a preoperational phase for geo-hazard monitoring and prevention (Esch et al. 2017). The CASEarth Poles comprise a comprehensive big data platform of the three poles including the Arctic, Antarctic and the Tibetan plateau within the framework of the "Big Earth Data Science and Engineering" program of the Chinese Academy of Science (Guo et al. 2017). One of the current initiatives is the NSF EarthCube originated from the NSF GEO Vision report (NSF Advisory Committee for Geosciences 2009). In this section, we introduce the EarthCube project and a big data infrastructure, Data Cube, as two cases of big data and cloud computing in the context of Digital Earth.

9.6.1 EarthCube

NSF EarthCube involves (1) Building Blocks (BBs), to develop novel infrastructure capabilities and demonstrate their value in a science context; (2) Research Coordination Networks (RCNs), to engage the science community around joint goals; (3) Conceptual Designs (CDs), to develop broad architecture design and explore integrative systems; (4) Integrative Activities (IAs), to explore concepts for the design of an enterprise architecture, and (5) Data Infrastructures (DIs) to lay the groundwork for shared data. The EarthCube concept originated from the NSF GEO Vision report (NSF Advisory Committee for Geosciences 2009), which was issued by the Advisory Committee for NSF's Geosciences Directorate (GEO) and identified the future focus of the Earth science community as '*fostering a sustainable future through a better understanding of our complex and changing planet*.' To achieve the GEO vision, the GEO and Office of Cyberinfrastructure (OCI) jointly launched the EarthCube (NSF 2011) initiative as a driving engine to build a geospatial cyberinfrastructure (similar to a Digital Earth infrastructure, Yang et al. 2010) to (1) understand and forecast the behavior of a complex and evolving Earth system; (2) reduce vulnerability and sustain life; and (3) train the workforce of the future.

EarthCube (2012) is targeted at (1) transforming the conduct of data-enabled geoscience-related research, (2) creating effective community-driven cyberinfrastructure, (3) allowing for interoperable resource discovery and knowledge management, and (4) achieving interoperability and data integration across disciplines (Fig. 9.4).

In addition, EarthCube is evolving within a rapidly growing, diverse, and wide-ranging global environment. In addition to the collaboration within EarthCube, there are other contributing entities ranging from individual data sets and software applications to national and international cyberinfrastructure systems. The NSF has funded the development of EarthCube through individual EarthCube awards since 2013. In 2016, the NSF awarded 11 new EarthCube activities, for a total of 51 awards. A sampling of efforts in EarthCube that benefit from big data and cloud computing are introduced below.

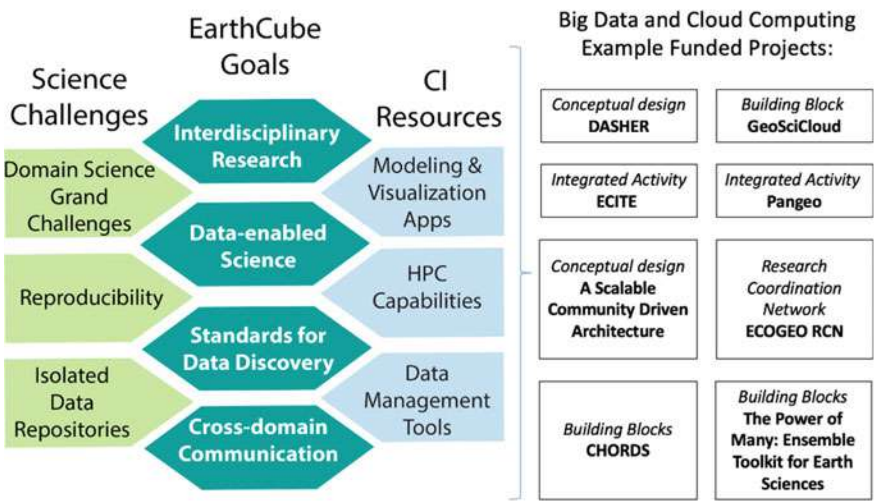


Fig. 9.4 Examples of related projects (derived from EarthCube goals, EarthCube Office 2016)

Conceptual design DASHER

Yang et al. (2015b) proposed a conceptual EarthCube Architecture, DASHER (Developing a Data-Oriented Human-Centric Enterprise Architecture for EarthCube), to support EarthCube and facilitate data communication and social collaboration in pursuit of collaborative Earth sciences research. The final product is a four-volume report containing different viewpoints that describe EarthCube architecture from different conceptual perspectives such as capabilities, operations, services, and projects. It provides a comprehensive conceptual reference for developing a detailed and practical architecture to address the requirements of the EarthCube community. DASHER was one of the first projects funded by EarthCube to design the conceptual framework integrating computational resources and big data sources.

Building Block GeoSciCloud

GeoSciCloud (Deploying Multi-Facility Cyberinfrastructure in Commercial and Private Cloud-based Systems) investigated two medium-size NSF funded data centers to deploy data collections with cloud-based services in different environments to assess feasibility and impact (EarthCube 2019). These environments include (1) commercial cloud environments offered by Amazon, Google, and Microsoft and (2) NSF-supported extensive computing facilities that are just beginning to offer services with characteristics of cloud computing.

GeoSciCloud helps EarthCube compare and contrast these three environments (the Extreme Science and Engineering Discovery Environment (XSEDE), commercial cloud, and current infrastructure) in the massive data ingestion to the cloud, data processing time, elasticity, the speed of data egress from multiple environments, overall costs of operation, interoperability, and reliability of real-time data streaming.

Integrated Activity ECITE

The EarthCube Integration and Test Environment (ECITE) is an outgrowth of activities of the EarthCube Testbed Working Group. The ECITE approach focuses on integrating existing effective technologies and resources as well as capabilities built by the EarthCube community using a cloud platform to provide a federated and interoperable test environment (EarthCube 2016). ECITE engages scientists and technologists from multiple disciplines and geographic regions across the Earth science community to develop requirements, prototype, design, build, and test an integration test-bed that will support cross-disciplinary research. The hybrid federated system will provide a robust set of distributed resources including both public and private cloud capabilities. This research addresses timely issues of integration, testing and evaluation methodologies and best practices with a strong interoperability theme to advance disciplinary research through the integration of diverse and heterogeneous data, algorithms, systems, and sciences.

Integrated Activity Pangeo

Pangeo[3] (an open-source big data climate science platform) integrates a suite of open-source software tools that can tackle petabyte-scale Atmosphere/Ocean/Land/Climate (AOC) datasets. Pangeo aims to cultivate an ecosystem in which the next generation of open-source analysis tools for ocean, atmosphere and climate science can be developed, distributed, and sustained. These tools must be scalable to meet the current and future challenges of big data, and the solutions should leverage the existing expertise outside of the AOC community. The resulting software improvements contribute to upstream open source projects, ensuring the long-term sustainability of the platform. The result is a robust new software toolkit for climate science and beyond. This toolkit will enhance the Data Science aspect of EarthCube. Implementation of these tools on the cloud was tested, taking advantage of an agreement between commercial cloud service providers and the NSF for big data solicitation.

9.6.2 Data Cube

The term 'data cube' was originally used in Online Analytical Processing (OLAP) of business and statistical data but has more recently been used in Earth domains as an approach to manage and analyze large and rapidly growing datasets. In Digital Earth, a data cube represents a multidimensional (n-D) array that stores gridded data or array-based data produced by remote sensing and simulation (Zhang et al. 2005). A data cube can be based on regular or irregular gridded, spatial and/or temporal data with multiple parameters. To support the management, sharing, and serving of Digital Earth data, tools and models, different cyberinfrastructures have been developed based on data cubes. Examples include the EarthServer that provides data cube services for Earth observations based on the RASDAMAN array database (Baumann et al. 2016). Another example is the Earth Observation Data and Processing Platform developed by the European Commission to integrate and analyze the combination of satellite and in situ Earth observations for sustainable development goals (Soille et al. 2016). The Committee on Earth Observation Satellites (CEOS) provides a data processing infrastructure based on data cubes to support Earth science objectives in developing countries, with a focus on remote sensing data (Nativi et al. 2017). The platform automatically ingests different remote sensing data into an N-dimensional data array.

Challenging issues in providing data services in data cube infrastructure include interoperability, rapid data access and transfer, and real-time processing and analysis (Strobl et al. 2017). Interoperability issues occur because datasets from various sources can have distinct parameterizations, spectral band definitions, projections, file formats, and database structures. One solution is to standardize the preprocessing procedure before storage and sharing with the community. The Open Geospatial

[3] http://pangeo.io/.

Consortium (OGC) Sensor Web Enablement (SWE) Common Data Model (CDM) defines important element parameterization (Robin 2011). Datasets must be represented along different predefined dimensions of the data cube, including space, time, and parameter properties. For projection difference issues, OGC recently developed the Discrete Global Grid System (DGGS) to optimize the loss of geospatial data during the reprojection process and seamlessly integrate GIS data from various sources (Stefanakis 2016). Data cube infrastructures also require rapid data access and transfer and real-time processing and analysis. Functionalities for user interactions must be built for various user demands, including file manipulation, data preprocessing, and analysis. These functionalities should also meet the standards of geographical information processing in OGC Web Coverage Services (WCS), geographic information analysis in the OGC Web Coverage Processing Service (WCPS), and format-independent data cube exchange in the OGC Coverage Implementation Schema (CIS).

Cloud computing and big data frameworks could enhance the data cube archive, visualization, and analysis in many ways to meet the needs of big Earth data knowledge mining. In cloud computing, storage is a virtual resource that can be attached and scaled on demand. By leveraging cloud computing and big data frameworks, visualizing data cubes and performing complicated spatiotemporal queries are more effortless than ever before (Zhizhin et al. 2011). One example of data cube visualization in the Earth science domain is the EOD[4] (Earth Observation Data Cube), which enables advanced data access and retrieval capabilities for the European coverage of Landsat-8 and the global coverage of Sentinel2 data. It aims to improve the accessibility of Big Earth data and offers more than 100 TB of Atmosphere, Land and Ocean EO products, demonstrating satellite data in the context of a virtual globe. The ESDC[5] (Earth System Data Cube) is another example of climate data cube visualization and analysis that aims to develop an environment to tap into the full potential of the ESA's Earth observations and integrate with the Biosphere-Atmosphere Virtual Laboratory (BAVL) analysis environment. The use of cloud computing technologies in big data visualization enables a massive amount of end users to explore data online at the same time with very low latency.

Data cube partition and parallel query could be achieved by utilizing distributed big data frameworks, which are faster and easier than traditional noncluster methods. Pagani et al. combined the data cube concept with cloud computing to manage and analyze large Earth datasets and observed better outcomes than traditional file-based approach (2018). Open Data Cube[6] is another example of the utilization of advances in cloud computing, providing free and open technologies to end users without local infrastructure. Thus, developing countries can access data and computing resources to build applications that aid decision making. The Australian Geoscience Data Cube (AGDA) solves similar problems of data sharing. It makes more than three decades of satellite imagery available for the first time, spanning Australia's total land area at

[4] https://eodatacube.eu/.

[5] https://cablab.readthedocs.io.

[6] https://www.opendatacube.org/.

a resolution of 25 square meters with more than 240,000 images that show how Australia's vegetation, land use, water movements, and urban expansion have changed over the past 30 years (NCI).

9.7 Conclusion

The advancement of Digital Earth drives collection of massive data such as transportation data, utility data, hazard data and forest data to monitor the Earth and support decision making. Valuable information extracted from the collected big data can further speed the development of Digital Earth. Digital Earth data continue to grow at a faster speed and with more heterogeneous types, leading to challenges to the lifecycle of data management including storage, processing, analytics, visualization, sharing, and integration. Fortunately, the emerging paradigm of cloud computing brings potential solutions to address these challenges. Compared with traditional computing mechanisms, cloud computing has the advantages of better data processing computing supports. The customizable configuration saves computing resources elastically, and data manipulation with higher security and flexibility offers secure data storage, transfer and sharing. Analytics enabled by cloud computing advance the process by allowing for automatic resource expansion when there are higher requirements.

To manage and analyze big Earth data, a service-oriented, scalable architecture based on cloud computing was introduced in a three-layer architecture: (1) the bottom data storage layer provides physical infrastructure, storage, and file systems; (2) the data query layer supplies data discovery capabilities with proper functionality and interoperability; and (3) the data processing layer supports extensibility, interoperability and scalability based on open source solutions and their variants from Earth science communities. With this architecture, big Earth data can be accessed and analyzed with low time latency or even in real time. The analysis results could be published by a web-based map server (e.g., GeoServer) or web-based notebook (e.g., Zeppelin) for visualization, public access, and collaboration, contributing to advancements in handling big data in Digital Earth to fulfill the requirements of scalability, extensibility and flexibility.

References

Agrawal R, Imieliński T, Swami A (1993) Mining association rules between sets of items in large databases. Acm SIGMOD Rec 22(2):207–216

Amirebrahimi S, Rajabifard A, Mendis P et al (2016) A framework for a microscale flood damage assessment and visualization for a building using BIM–GIS integration. Int J Digit Earth 9(4):363–386

Anderson A (2015) Statistics for big data for dummies. John Wiley & Sons, Hoboken, NJ

Balakrishna C (2012) Enabling technologies for smart city services and applications. In: 2012 sixth International conference on next generation mobile applications, services and technologies. IEEE, Paris, France, 12–14 September 2012
Baumann P, Mazzetti P, Ungar J et al (2016) Big data analytics for earth sciences: the earthserver approach. Int J Digit Earth 9(1):3–29
Bereta K, Caumont H, Daniels U et al (2019) The copernicus app lab project: easy access to copernicus data. In: EDBT. pp 501–511
Big Data Virtualization (2019) https://www.techopedia.com/definition/29952/big-data-virtualization. Accessed 6 May 2019
Bizer C, Heath T, Berners-Lee T (2011) Linked data: the story so far. In: Amit S (ed) Semantic services, interoperability and web applications: emerging concepts. IGI Global, Hershey, PA, pp 205–227
Blachowski J (2016) Application of GIS spatial regression methods in assessment of land subsidence in complicated mining conditions: case study of the Walbrzych coal mine (SW Poland). Nat Hazards 84(2):997–1014
Boulos MNK, Al-Shorbaji NM (2014) On the internet of things, smart cities and the WHO Healthy Cities. Int J Health Geogr 13:10
Buck JB, Watkins N, LeFevre J et al (2011) SciHadoop: array-based query processing in Hadoop. In: Proceedings of 2011 International conference for high performance computing, networking, storage and analysis, ACM, New York, NY, 12–18 Nov 2011
Canada Line Vancouver Transit Map (2019) https://airfreshener.club/quotes/canada-line-vancouver-transit-map.html. Accessed 6 May 2019
Chang WL, Grady N (2015) NIST big data interoperability framework: volume 1, big data definitions (No. special publication (NIST SP)-1500-1).
Cudre-Mauroux P, Kimura H, Lim K-T et al (2009) A demonstration of SciDB: a science-oriented DBMS. Proc VLDB Endow 2(2):1534–1537
Dempsey C (2012) Where is the phrase "80% of data is geographic" from. https://www.gislounge.com/80-percent-data-is-geographic. Accessed 6 May 2019
Dillon T, Wu C, Chang E (2010) Cloud computing: issues and challenges. In: 2010 24th IEEE international conference on advanced information networking and applications, IEEE, Perth, Western Australia, 20–23 Apr 2010
Duffy DQ, Schnase JL, Thompson JH et al (2012) Preliminary evaluation of mapreduce for high-performance climate data analysis. NASA new technology report white paper
EarthCube (2016) EarthCube integration and testing environment(ECITE). https://www.earthcube.org/group/earthcube-integration-testing-environment-ecite. Accessed 6 May, 2019
EarthCube (2019) GeoSciCloud: Deploying Multi-Facility Cyberinfrastructure in Commercial and Private Cloud-based Systems https://www.earthcube.org/group/geoscicloud-deploying-multi-facility-cyberinfrastructure-commercial-private-cloud-based-systems. Accessed 6 May, 2019
EarthCube Brochure (2012) What is EarthCube? http://www.azgs.az.gov/images/agu-2012-earthcube-brochure-1.pdf. Accessed 11 Dec 2018
Eldawy A, Mokbel MF (2015) Spatial hadoop: a mapreduce framework for spatial data. In: 2015 IEEE 31st international conference on data engineering, IEEE, Seoul, South Korea, 13–17 Apr 2015
El-Mekawy M (2010) Integrating BIM and GIS for 3D city modelling: the case of IFC and CityGM. Doctoral Dissertation, KTH
Esch T, Uereyen S, Asamer H et al (2017) Earth observation-supported service platform for the development and provision of thematic information on the built environment—the TEP-Urban project. In: 2017 joint urban remote sensing event (JURSE), IEEE, Dubai, UAE, 6–8 Mar 2017
Firican G (2017) The 10 Vs of big data. Upside where data means business. https://tdwi.org/articles/2017/02/08/10-vs-of-big-data.aspx. Accessed 30 Jul 2019
Foster I, Zhao Y, Raicu I et al (2008) Cloud computing and grid computing 360-degree compared. arXiv preprint arXiv:0901.0131.

Friedman U (2012) Big data: a short history. https://foreignpolicy.com/2012/10/08/big-data-a-short-history. Accessed 6 May 2019
Gantz J, Reinsel D (2011) Extracting value from chaos. IDC iview 1142(2011):1–12
Geng Y, Huang X, Zhu M et al (2013) SciHive: array-based query processing with HiveQL. In: 2013 12th IEEE international conference on trust, security and privacy in computing and communications, IEEE, Melbourne, Australia, 16–18 Jul 2013
Geng Y, Huang X, Yang G (2014) Adaptive indexing for distributed array processing. In: 2014 IEEE international congress on big data, IEEE, Anchorage, AK, 27 June-2 Jul 2014
Giuliani G, Lacroix P, Guigoz Y et al (2017) Bringing GEOSS services into practice: a capacity building resource on spatial data infrastructures (SDI). Trans GIS 21(4):811–824
Gundersen E (2013) Visualizing 3 billion tweets. https://blog.mapbox.com/visualizing-3-billion-tweets-f6fc2aea03b0. Accessed 6 May 2019
Guo H (2017) Big earth data: a new frontier in earth and information sciences. Big Earth Data 1(1–2):4–20
Guo Z, Feng C-C (2018) Using multi-scale and hierarchical deep convolutional features for 3D semantic classification of TLS point clouds. Int J Geogr Inform Sci 1–20. https://doi.org/10.1080/13658816.2018.1552790
Guo H, Liu Z, Jiang H et al (2017) Big earth data: a new challenge and opportunity for digital earth's development. Int J Digit Earth 10(1):1–12
Hacker H, Trinitis C, Weidendorfer J et al (2010) Considering GPGPU for HPC centers: is it worth the effort? In: Keller R, Kramer D, Weiss J-P (eds) Facing the multicore-challenge: aspects of new paradigms and technologies in parallel computing. Springer, Berlin, Heidelberg, pp 118–130
Hamscher V, Schwiegelshohn U, Streit A et al (2000) Evaluation of job-scheduling strategies for grid computing. In: Buyya R, Baker M (eds) Grid computing—GRID 2000. Springer, Berlin, Heidelberg, pp 191–202
Hashem IAT, Chang V, Anuar NB et al (2016) The role of big data in smart city. Int J Inform Manag 36(5):748–758
Heuvelink GBM, Pebesma EJ (1999) Spatial aggregation and soil process modelling. Geoderma 89(1):47–65
Hong S-Y, O'Sullivan D (2012) Detecting ethnic residential clusters using an optimisation clustering method. Int J Geogr Inform Sci 26(8):1457–1477
Hu F, Xu M, Yang J et al (2018a) Evaluating the open source data containers for handling big geospatial raster data. ISPRS Int J Geoinform 7(4):144
Hu F, Yang C, Schnase JL et al (2018b) ClimateSpark: an in-memory distributed computing framework for big climate data analytics. Comput Geosci 115:154–166
Huffman GJ, Bolvin DT, Braithwaite D et al (2015) NASA global precipitation measurement (GPM) integrated multi-satellite retrievals for GPM (IMERG). In: Algorithm theoretical basis document, version 4.1. NASA, Washington, DC
Hughes JN, Annex A, Eichelberger CN et al (2015) Geomesa: a distributed architecture for spatio-temporal fusion. In: Proceedings of SPIE 9473, geospatial informatics, fusion, and motion video analytics, SPIE, Washington, DC, 21 May 2015
Jiang Y, Li Y, Yang C et al (2018b) Towards intelligent geospatial data discovery: a machine learning framework for search ranking. Int J Digit Earth 11(9):956–971
Jiang Y, Li Y, Yang C et al (2018a) A smart web-based geospatial data discovery system with oceanographic data as an example. ISPRS Int J Geoinform 7(2):62
JoSEP, A. D., KAtz, R., KonWinSKi, A., Gunho, L. E. E., PAttERSon, D., & RABKin, A. (2010). A view of cloud computing. Communications of the ACM, 53(4).
JPL (2001) Izmit, Turkey 1999 earthquake interferogram. https://www.jpl.nasa.gov/spaceimages/details.php?id=PIA00557. Accessed 6 May 2019
Kofinas DT, Spyropoulou A, Laspidou CS (2018) A methodology for synthetic household water consumption data generation. Environ Model Softw 100:48–66
Krämer M, Senner I (2015) A modular software architecture for processing of big geospatial data in the cloud. Comput Graphics 49:69–81

Lahsen M (2016) Toward a sustainable future earth: challenges for a research agenda. Sci Technol Hum Values 41(5):876–898
Lam NSN (1983) Spatial interpolation methods: a review. Am Cartogr 10(2):129–150
LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444
Lenka RK, Barik RK, Gupta N et al (2016) Comparative analysis of SpatialHadoop and GeoSpark for geospatial big data analytics. In: 2016 2nd international conference on contemporary computing and informatics (IC3I), IEEE, Noida, India, 14–17 Dec 2016
Lepot M, Aubin J-B, Clemens HF (2017) Interpolation in time series: an introductive overview of existing methods, their performance criteria and uncertainty assessment. Water 9(10):796
Li W, Hsu C-Y (2018) Automated terrain feature identification from remote sensing imagery: a deep learning approach. Int J Geogr Inform Sci 1–24. https://doi.org/10.1080/13658816.2018.1542697
Li A, Yang X, Kandula S et al (2010) CloudCmp: comparing public cloud providers. In: Proceedings of the 10th ACM SIGCOMM conference on internet measurement, ACM, Melbourne, Australia, 1–30 Nov 2010
Li C, Wang S, Kang L et al (2014a) Trust evaluation model of cloud manufacturing service platform. Int J Adv Manuf Technol 75(1):489–501
Li W, Goodchild MF, Raskin R (2014b) Towards geospatial semantic search: exploiting latent semantic relations in geospatial data. Int J Digit Earth 7(1):17–37
Li Z, Yang C, Jin B et al (2015) Enabling big geoscience data analytics with a cloud-based, MapReduce-enabled and service-oriented workflow framework. PLoS ONE 10(3):e0116781
Li S, Dragicevic S, Castro FA et al (2016) Geospatial big data handling theory and methods: a review and research challenges. ISPRS J Photogramm Remote Sens 115:119–133
Li Z, Hu F, Schnase JL et al (2017a) A spatiotemporal indexing approach for efficient processing of big array-based climate data with MapReduce. Int J Geogr Inform Sci 31(1):17–35
Li Z, Yang C, Huang Q et al (2017b) Building model as a service to support geosciences. Comput Environ Urb Syst 61:141–152
Lipponen A (2017) New year's eve – average temperature at 18:00 local time. https://www.flickr.com/photos/150411108@N06/38663730944. Accessed 6 May 2019
Liu K, Yang C, Li W et al (2011) The GEOSS clearinghouse high performance search engine. In: 2011 19th international conference on geoinformatics, IEEE, Shanghai, China, 24–26 June 2011
Lui K, Yang C, Gui Z (2013) GeoSearch: a system utilizing ontology and knowledge reasoning to support geospatial data discovery. In: Workshop on semantics in geospatial architectures: applications and implementation, Pyle Center, University of Wisconsin-Madison, Madison, Wisconsin, 28–29 Oct 2013
Malik T (2014) GeoBase: indexing NetCDF files for large-scale data analysis. In: Wen-Chen H, Naima K (eds) Big data management, technologies, and applications. IGI Global, Hershey, PA, pp 295–313
Marr B (2015) Big data: using SMART big data, analytics and metrics to make better decisions and improve performance. John Wiley & Sons, Hoboken, NJ
Mayama K, Skulkittiyut W, Ando Y et al (2011) Proposal of object management system for applying to existing object storage furniture. In: 2011 IEEE/SICE international symposium on system integration (SII), IEEE, Kyoto, Japan, 20–22 Dec 2011
McCandless M, Hatcher E, Gospodnetic O (2010) Lucene in action: covers apache lucene 3.0. Manning Publications Co., New York, NY
Mc Cutchan M (2017) Linked data for a digital earth: spatial forecasting with next generation geographical data. In: International conference on spatial information theory, Springer, Cham, pp 91–96
Mell PM, Grance T (2011) Sp 800–145. The nist definition of cloud computing. NIST, Gaithersburg, MD
Microsoft (2011) Microsoft researchers' focus: eye on earth. https://www.microsoft.com/en-us/research/blog/microsoft-researchers-focus-eye-on-earth. Accessed 6 May 2019

NASA (2016) Getting petabytes to people: How the EOSDIS facilitates earth observing data discovery and use. https://earthdata.nasa.gov/getting-petabytes-to-people-how-the-eosdis-facilitates-earth-observing-data-discovery-and-use. Accessed 6 May 2019

Nativi S, Mazzetti P, Santoro M et al (2015) Big data challenges in building the global earth observation system of systems. Environ Model Softw 68:1–26

Nativi S, Mazzetti P, Craglia M (2017) A view-based model of data-cube to support big earth data systems interoperability. Big Earth Data 1(1–2):75–99

NCI (2019) Australian geoscience data cube. http://nci.org.au/services/virtual-laboratories/australian-geoscience-data-cube. Accessed 6 May 2019

NIST Big Data Public Working Group (2015) Nist big data interoperability framework. Use cases and general requirements. NIST, Maryland, US

NOAA (2017) GOES-16 first SEISS data.png. https://commons.wikimedia.org/wiki/File:GOES-16_first_SEISS_data.png. Accessed 6 May 2019

NSF (2009) Advisory committee for geosciences. http://www.nsf.gov/geo/acgeo/geovision/nsf_ac-geo_vision_10_2009.pdf. Accessed 11 Dec 2018

NSF (2011) Earth cube guidance for the community. http://www.nsf.gov/pubs/2011/nsf11085/nsf11085.pdf.. Accessed 11 Dec 2018

Pagani GA, Trani L (2018) Data cube and cloud resources as platform for seamless geospatial computation. In: Proceedings of the 15th ACM international conference on computing frontiers, ACM, Ischia, Italy, 8–10 May 2018

Qi H, Gani A (2012) Research on mobile cloud computing: review, trend and perspectives. In: 2012 second international conference on digital information and communication technology and it's applications (DICTAP), IEEE, Bangkok, Thailand, 16–18 May 2012

Radke AM, Tseng MM (2015) Design considerations for building distributed supply chain management systems based on cloud computing. J Manuf Sci Eng 137(4):040906

Rahman MR, Lateh H (2017) Climate change in Bangladesh: a spatio-temporal analysis and simulation of recent temperature and rainfall data using GIS and time series analysis model. Theor Appl Climatol 128(1):27–41

Rathore P, Rao AS, Rajasegarar S et al (2017) Real-time urban microclimate analysis using internet of things. IEEE Internet Things J 5(2):500–511

Rew R, Davis G (1990) NetCDF: an interface for scientific data access. IEEE Comput Graph Appl 10(4):76–82

Rienecker MM, Suarez MJ, Gelaro R et al (2011) MERRA: NASA's modern-era retrospective analysis for research and applications. J Clim 24(14):3624–3648

Robert K (2000) Data flow controls water flow. https://www.spk.usace.army.mil/Media/Images/igphoto/2000748857. Accessed 6 May 2019

Robin A (2011) SWE CDM encoding standard, OGC. http://www.opengeospatial.org/standards/swecommon. Accessed 6 May 2019

Schabenberger O, Gotway CA (2017) Statistical methods for spatial data analysis. CRC Press, Boca Raton, FL

Segal B, Robertson L, Gagliardi F et al (2000) Grid computing: the European data grid project. In: 2000 IEEE nuclear science symposium. Conference record (Cat. No.00CH37149). IEEE, Lyon, France, 15–20 Oct 2000

Sharifzadeh M, Shahabi C (2004) Supporting spatial aggregation in sensor network databases. In: Proceedings of the 12th annual ACM international workshop on geographic information systems, ACM, New York, NY, 12–13 Nov 2004

Shi W, Cao J, Zhang Q et al (2016) Edge computing: vision and challenges. IEEE Internet Things J 3(5):637–646

Söderberg A, Dahlström P (2017) Turning smart water meter data into useful information: a case study on rental apartments in södertälje. Lund University, Lund, Sweden

Soille P, Burger A, Rodriguez D et al (2016) Towards a JRC earth observation data and processing platform. In: Proceedings of the conference on big data from space (BiDS'16), Publications Office of the European Union, Santa Cruz de Tenerife, 15–17 Mar 2016

Stefanakis E (2016) Discrete global grid systems—a new OGC standard emerges. GoGeomatics: Magazine of Gogeomatics Canada

Strobl P, Baumann P, Lewis A et al (2017) The six faces of the data cube. In: Proceedings of conference on big data from space (BiDS'17), Toulouse, France, 28–30 Nov 2017

USGS (2019) What is remote sensing and what is it used for? https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used?qt-news_science_products=7#qt-news_science_products. Accessed 6 May 2019

Vilches-Blázquez LM, Villazón-Terrazas B, Corcho O et al (2014) Integrating geographical information in the linked digital earth. Int J Digit Earth 7(7):554–575

Wang S, Liu Y (2009) Teragrid giscience gateway: bridging cyberinfrastructure and giscience. Int J Geographi Inf Sci 23(5):631–656

White T (2012) Hadoop: the definitive guide. O'Reilly Media, Inc., Sebastopol, CA

Xu C, Yang C, Li J et al (2011) A service visualization tool for spatial web portal. In: Proceedings of the 2nd international conference on computing for geospatial research & applications, ACM, New York, NY, 23–25 May 2011

Yang R (2016) A systematic classification investigation of rapid intensification of atlantic tropical cyclones with the ships database. Weather Forecast 31(2):495–513

Yang C, Raskin R, Goodchild M et al (2010) Geospatial cyberinfrastructure: past, present and future. Comp Environ Urban Sys 34(4):264–277

Yang C, Goodchild M, Huang Q et al (2011a) Spatial cloud computing: how can the geospatial sciences use and help shape cloud computing? Int J Digit Earth 4(4):305–329

Yang R, Tang J, Sun D (2011b) Association rule data mining applications for atlantic tropical cyclone intensity changes. Weather Forecast 26(3):337–353

Yang C, Sun M, Liu K et al (2015a) Contemporary computing technologies for processing big spatiotemporal data. In: Kwan M-P, Richardson D, Wang D et al (eds) Space-time integration in geography and GIScience: research frontiers in the US and China. Springer Netherlands, Dordrecht, pp 327–351

Yang CP, Yu M, Sun M et al (2015b) Dasher cd: developing a data-oriented human-centric enterprise architecture for earthcube. In: AGU fall meeting abstracts. AGU, Washington, DC

Yang C, Huang Q, Li Z et al (2017b) Big data and cloud computing: innovation opportunities and challenges. Int J Digital Earth 10(1):13–53

Yang C, Yu M, Hu F et al (2017a) Utilizing cloud computing to address big geospatial data challenges. Comp EnvironUrban Syst 61:120–128

Yoo C, Ramirez L, Liuzzi J (2014) Big data analysis using modern statistical and machine learning methods in medicine. Int Neurourol J 18(2):50–57

Yu J, Wu J, Sarwat M (2015) Geospark: a cluster computing framework for processing large-scale spatial data. In: Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems, ACM, Seattle, US, 3–6 Nov 2015

Yu M, Yang C, Li Y (2018) Big data in natural disaster management: a review. Geosciences 8(5):165

Zaharia M, Chowdhury M, Das T et al (2012) Resilient distributed datasets: a fault-tolerant abstraction for in-memory cluster computing. In: Proceedings of the 9th USENIX conference on networked systems design and implementation, USENIX Association, Berkeley, CA, 25–27 Apr 2012

Zhang Y, Kerle N (2008) Satellite remote sensing for near-real time data collection. In: Geospatial information technology for emergency response, 1st edn. CRC Press, Boca Raton, FL, pp 91–118

Zhang Y, Kunqing X, Xiujun M et al (2005) Spatial data cube: provides better support for spatial data mining. In: Proceedings 2005 IEEE international geoscience and remote sensing symposium. IGARSS'05, IEEE, Seoul, South Korea, 29–29 Jul 2005

Zhao H, Ai S, Lv Z et al (2010) Parallel accessing massive NetCDF data based on MapReduce. In: Wang FL, Gong Z, Luo X, Lei J (eds) International Conference on web information systems and mining. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 425–431

Zhao P, Qin K, Ye X et al (2017) A trajectory clustering approach based on decision graph and data field for detecting hotspots. Int J Geographi Inf Sci 31(6):1101–1127

Zhizhin M, Medvedev D, Mishin D et al (2011) Transparent data cube for spatiotemporal data mining and visualization. In: Fiore S, Aloisio G (eds) Grid and cloud database management. Springer, Berlin, Heidelberg, pp 307–330

Zikopoulos B, Barbas H. (2012). Pathways for emotions and attention converge on the thalamic reticular nucleus in primates. Journal of Neuroscience, 32(15), 5338–5350.

Yun Li is a Ph.D. candidate majoring in Earth Systems and Geoinformation Sciences at George Mason University. Her research mainly focuses on improving Earth data discovery using machine learning based methods, high performance computing, and spatiotemporal analytics for environmental and climate data.

Manzhu Yu is an Assistant Professor of GIScience at Pennsylvania State University. Her research focuses on spatiotemporal theories and applications, atmospheric modeling, environmental analytics, big data and cloud computing, and the ability of using the above to solve pressing issues in natural hazards and sustainability.

Mengchao Xu is a Ph.D. candidate in Earth Systems and Geoinformation Sciences program from George Mason University. His research mainly focuses on the cloud computing, high performance computing networks, distributed database systems, and precipitation downscaling. He manages a 500-nodes cluster specialized for cloud computing (supported by NASA, NSF and GMU).

Jingchao Yang is a Ph.D. student majoring in Earth Systems and Geoinformation Sciences at George Mason University. He has been contributing to several NSF and NASA funded projects at the NSF Spatiotemporal Innovation Center. His research mainly focuses on the smart city problems with machine- and deep-learning enabled spatiotemporal pattern analysis using multi-sensor (IoT) datasets.

Dexuan Sha is a Ph.D. student in Earth Systems and Geoinformation Sciences at the George Mason University. He researches arctic sea ice and melt pond from high spatial resolution aerial imagery with big spatial–temporal data mining and analysis methodology.

Qian Liu is a Ph.D. student in Geography and Geoinformation Science at George Mason University. Her research mainly focuses on the core areas of precipitation detection and analysis using AI-based methods, and outreaches to environmental and climate science and spatial temporal analytics.

Chaowei Yang is Professor of GIScience at George Mason University, where he established and directs the NSF spatiotemporal innovation center. His academic activities include spatiotemporal computing research, future GIScience leader education, and Geoinformatics outreach.

Chapter 10
Artificial Intelligence

Eric Guérin, Orhun Aydin and Ali Mahdavi-Amiri

Abstract In this chapter, we provide an overview of different artificial intelligence (AI) and machine learning (ML) techniques and discuss how these techniques have been employed in managing geospatial data sets as they pertain to Digital Earth. We introduce statistical ML methods that are frequently used in spatial problems and their applications. We discuss generative models, one of the hottest topics in ML, to illustrate the possibility of generating new data sets that can be used to train data analysis methods or to create new possibilities for Digital Earth such as virtual reality or augmented reality. We finish the chapter with a discussion of deep learning methods that have high predictive power and have shown great promise in data analysis of geospatial data sets provided by Digital Earth.

Keywords Artificial intelligence · Machine learning · Generative models · Statistical data analysis

10.1 Introduction

Earth and its associated data sets are massive. Various forms of geospatial data sets are constantly accumulated and captured by different forms of sensors and devices (Mahdavi-Amiri et al. 2015). Managing such an immense data set is a challenge. As a result, many automated techniques have been designed to process geospatial data sets with minimal human interference. Since manual involvement should be minimal, the machines should be capable of processing data and delivering meaningful information to the users. With advancements in machine learning, processing

E. Guérin
LIRIS, Université de Lyon, Lyon, France

O. Aydin
ESRI, Los Angeles, Los Angeles, USA

University of Southern California, Los Angeles, USA

A. Mahdavi-Amiri (✉)
Department of Computing Science, Simon Fraser University, Burnaby, Canada
e-mail: amahdavi@sfu.ca

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_10

geospatial data sets has significantly improved. In this chapter, we discuss artificial intelligence and machine learning techniques that have been useful to manage and process geospatial data sets. Because the processing of geospatial data can also be a source of knowledge, some methods use existing data to generate and synthesize new data.

We start by discussing some traditional and statistical approaches in machine learning and then present more recent learning techniques employed for geospatial data sets. Traditional methods include predefined models such as linear regression, PCA, SVD, active contour, and SVM, in which the model is fixed and the learning is based on an optimization. We also briefly discuss evolutionary and agent-based methods and autoencoders as traditional methods that can be deep or shallow. We then discuss more recent deep learning techniques, including reinforcement learning, deep convolutional networks and generative models such as variational autoencoders and generative adversarial networks. In this chapter, we describe some applications of these machine learning techniques to handle geospatial data sets that are the main content of Digital Earth. In the future, a dynamic Digital Earth that can use such techniques to work with geospatial datasets is extremely practical. Currently, such methods are sparsely used on very specific Digital Earth data sets. We imagine that a more advanced Digital Earth will use state-of-the-art machine learning techniques much more than they are currently used.

10.2 Traditional and Statistical Machine Learning

Inferring patterns and forming relationships using artificial intelligence require knowledge of some characteristics of the phenomena/system of interest. One of the early approaches to enabling artificial intelligence for complex problems was to create knowledge bases that contain explicit sets of rules and associations, also known as ontology (Gruber 1993). For data pertaining to Earth system modeling, different niche knowledge bases were designed by various authors (McCarthy 1988; Rizzoli and Young 1997). The knowledge base approach to artificial intelligence required expert input to define the rules and associations. In addition, the expert knowledge had to be represented in a "computable form" (Sowa 2000), posing a bottleneck for these approaches. For spatially varying, complex phenomena, ontology representations were defined for Earth's subsystems such as in environmental modeling and planning (Cortés et al. 2001), and ecological reasoning (Rykiel 1989). General spatial and GIS knowledge bases were proposed by various authors (Kuipers 1996; Egenhofer and Mark 1995; Fonseca et al. 2002).

Despite the plethora of niche knowledge bases, knowledge base artificial intelligence requires assertions and ground truths (Lenat 1995), which can conflict with observations (Goodfellow et al. 2016). Numerous attempts to address this limitation have been presented by various authors, such as defining hierarchical (Kuipers 1996), or location/problem-tailored knowledge bases (Rizzoli and Young 1997).

Statistical machine learning alleviates the limitations of the knowledge-based approach to artificial intelligence and discovers rules and patterns from the data directly without explicit supervision (Goodfellow et al. 2016). In the case of statistical learning, patterns and rules from an unknown underlying process are defined for descriptive, predictive and prescriptive analytics.

Applications of statistical learning to understand and forecast natural and human phenomena are evaluated with respect to the components of the general definition of machine learning (Mitchell 1997). Mitchell's (1997) definition is as follows:

> A computer program is said to learn from experience [D] with respect to some class of tasks T and performance measure [Q], if its performance at tasks in T, as measured by [Q], improves with experience [D].

Machine learning methods are broadly grouped into supervised and unsupervised methods. Supervised machine learning methods experience modeled phenomena through so-called labeled training data. Labels in the training data correspond to the target variable to be predicted, either quantitative (regression) or qualitative (classification). Training data consists of predictors and their corresponding predictand. Thus, supervised machine learning methods learn relationships in the data through experiencing input/output pairs.

Unsupervised machine learning methods discover patterns in the data without supervision or explicit rules. Clustering is one of the most common unsupervised machine learning methods for geospatial datasets.

10.2.1 Supervised Learning

Supervised learning aims to define a relationship between r predictor variables, denoted by $X = (X_1, X_2, \ldots, X_r)$, and e predictands, $Y = (Y_1, Y_2, \ldots, Y_e)$. Supervised learning can be posed as a density estimation problem (Hastie et al. 2001):

$$\boldsymbol{P}(\boldsymbol{Y}|\boldsymbol{X}) = \boldsymbol{P}(\boldsymbol{Y}, \boldsymbol{X})/\boldsymbol{P}(\boldsymbol{X}) \tag{10.1}$$

where $P(Y|X)$ is the conditional probability density of observing the predictand given the predictors, $P(Y, X)$ is the joint probability distribution of the predictand and predictors, and $P(X)$ is the marginal probability distribution of the predictors. Using Mitchell's (1997) description, the performance Q can be quantified using a loss function $\mathcal{L}$ where, for a given method and set of parameters Θ, a location function, $\mu(x)$, is minimized (Hastie et al. 2001) in Eq. 10.2.

$$\boldsymbol{\mu}(\boldsymbol{x}) = \boldsymbol{argmin}_{\boldsymbol{\Theta}} \boldsymbol{E}_{\boldsymbol{Y}|\boldsymbol{X}} \boldsymbol{\mathcal{L}}(\boldsymbol{Y}, \boldsymbol{\Theta}) \tag{10.2}$$

For a given Θ, a supervised machine learning method predicts the values at X as $\hat{y}$. The loss function, $\mathcal{L}$, quantifies the error between $\hat{y}$ and the training data y.

Some examples of supervised machine learning methods as they pertain to geospatial analysis are given in the following subsection.

10.2.1.1 Random Forest

Random forest is a framework for nonparametric estimation in which both classification and regression can be performed (Breiman 2001). It has gained popularity in numerous geospatial applications due to its flexibility in accommodating different types of inputs (categorical or continuous) and its ability to model complex relationships in the data.

Random forest addresses the overfitting limitation of **c**lassification **a**nd **r**egression **t**rees (CART). Random forest uses bootstrap aggregating, also known as bagging, to create subsets of the training data by sampling with replacement to build different CARTs (Breiman 1996). Each of the CARTs that make up the forest predict, or vote, for a given data point of x and the forest returns the majority vote in a classification or the average forest prediction for a regression. The voting scheme of random forest allows for complex relationships to be captured in the data that might not be possible otherwise. A pictorial summary of a random forest classifier for classifying a successful retail store (one) or an unsuccessful one (zero) with respect to its distance to the nearest highway exit and the number of brands it carries is given in Fig. 10.1.

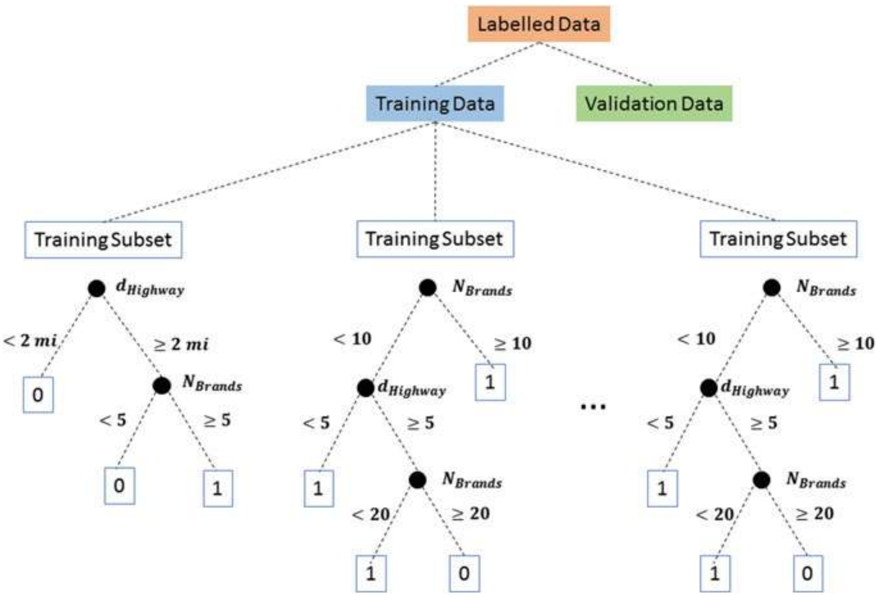


Fig. 10.1 Cartoon representation of a random forest classifier

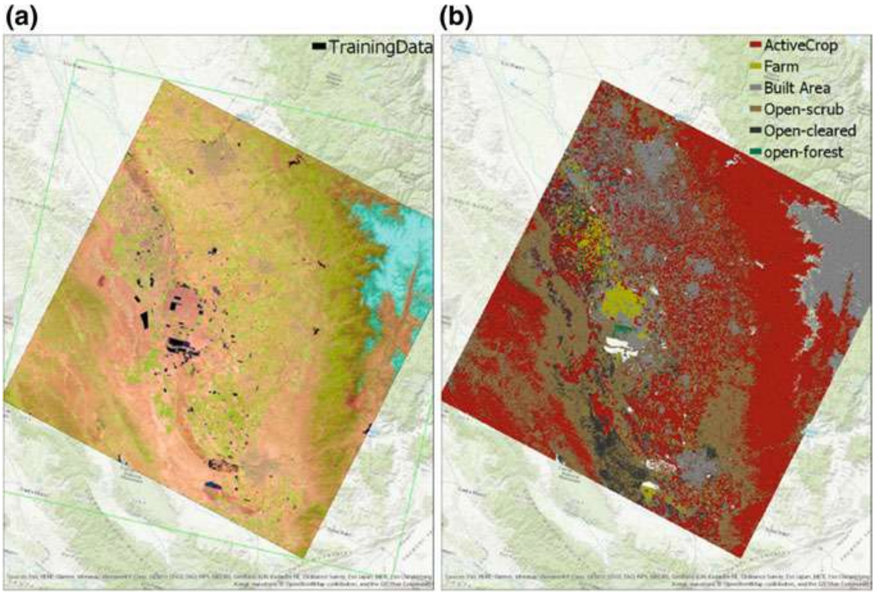


Fig. 10.2 **a** Satellite image over southern California, with training data marked with black polygons **b** classified land coverage map using random forest

Note that every tree experiences different subsets of training data and their structures are different from one another. The voting scheme allows for capturing underlying patterns in the data by defining complex relationships captured in a large ensemble of trees rather than a single tree.

In geospatial problems, various random forest classifiers are used in a wide range of problems, including land cover classification (Gislason et al. 2006) and ecological modeling (Cutler et al. 2007). In land cover classification, random forest speeds up classification of land use by forming a relationship between the satellite image RGB value and the type of land it corresponds to. In this case, the training data consists of tagged locations at which the land cover and RGB values are known. An example of the random forest classifier output for land use classification is given in Fig. 10.2.

In Fig. 10.2, a small number of farms and areas around them were used as training data (marked with black polygons). The training set that consists of 300 farms was used within the random forest classifier to define land use in southern California.

10.2.1.2 Geographically Weighted Regression

Geographically **w**eighted **r**egression (GWR) provides a statistical framework for incorporating spatial dependency within a linear regression system (Fotheringham

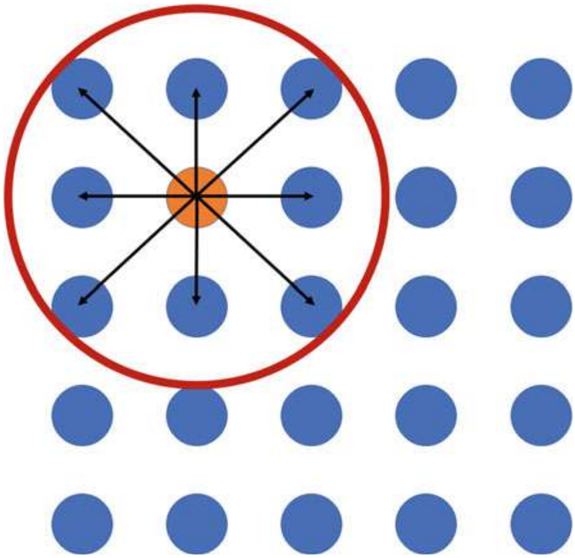


Fig. 10.3 Conceptual depiction of GWR. Regression is performed for the orange point with a red circle defining the neighborhood

et al. 2003). GWR provides spatial extensions to ordinary least squares and generalized linear models (Nelder and Wedderburn 1972) such as geographically weighted logistic regression. GWR is depicted conceptually in Fig. 10.3.

Figure 10.3 illustrates a regression system solved within the neighborhood (red circle) for the location indicated in orange. First, GWR defines a weighting scheme to determine spatial weights for the neighbors, and the predictors $\boldsymbol{X}$ at every location (blue) are weighted with respect to their distance to the location for which the regression is performed (orange). The geographically weighted linear system of equations solved at a point i can be expressed as follows:

$$\hat{\boldsymbol{\beta}}(\boldsymbol{u}_i, \boldsymbol{v}_i) = \left(\mathbf{X}^T \mathbf{W}(\boldsymbol{u}_i, \boldsymbol{v}_i)\mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{W}(\boldsymbol{u}_i, \boldsymbol{v}_i)\mathbf{Y} \tag{10.3}$$

where $\hat{\boldsymbol{\beta}}(\boldsymbol{u}_i, \boldsymbol{v}_i)$ is the coefficient matrix for the predictors $\mathbf{X}$ at location i. $\mathbf{W}(u_i, v_i)$ is a diagonal weighting matrix that contains geographic weights on its diagonal elements for neighbors inside the neighborhood window (red circle in Fig. 10.3), and $\mathbf{Y}$ contains the variable being predicted. Note that the linear system above is similar to the general linear regression system given in Eq. 10.4.

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}^T \mathbf{X}\right)^{-1} \mathbf{X}^T \mathbf{Y} \tag{10.4}$$

where $\hat{\boldsymbol{\beta}}$ is defined globally for the entire dataset. The geographic weights are inversely weighted with respect to the distance. Thus, the weights have large values for neighbors close to the regression location i. Different weighting schemes and neighborhood definitions are possible; the reader is encouraged to explore seminal work on this topic (Fotheringham et al. 2003).

Spatial representation via a weighting scheme can give GWR high predictive power for geospatial datasets in which a strong spatial autocorrelation is observed. The impact of incorporating spatial relationships in the regression model is demonstrated by comparing GWR with a nonspatial supervised machine learning method. In this example, GWR is juxtaposed against a random forest predictor for a problem with strong spatial autocorrelation in the data. Statistical climate downscaling (Wilby and Wigley 1997) was performed with GWR and a random forest regressor. Statistical downscaling calibrates the output of a global circulation model (GCM) to observed climate data such as temperature or precipitation. In this example, climate downscaling for the lower 48 US states; a regression model can be defined between 19 predictors (from GCM) and the observed average temperature. The regression model can be used to predict the average temperature for the entire lower 48 states. A random forest predictor can be trained using the observed average temperature and simulated GCM variables. The GWR model is formed using only 3 of the independent predictors due to the collinearity restriction of GWR. Below are the predicted average temperature profiles.

Note that the average temperature profile estimated in Fig. 10.4a depicts the patterns of temperature change captured in Fig. 10.4b. Even though fewer predictors are used in the GWR than in the random forest regressor, large-scale patterns in the temperature profile changes are captured. The GWR model in Fig. 10.4a was also compared to a random forest regressor model trained using the same three predictors. In that case, the GWR returned a mean-squared error that was 60% of that of the random forest regressor.

10.2.1.3 SVM

Support vector machine (SVM) is a supervised nonparametric statistical learning method (Corinna and Vapnik 1995). In its original form, the method comprises a set of labeled data instances and the SVM attempts to find a hyperplane that separates the dataset into a discrete predefined number of classes as consistently as possible for the training data (see Fig. 10.5) (Vapnik 1979). It is possible to generalize SVM to nonlinear kernels such as radial basis functions to learn and classify data sets with higher complexity (Schölkopf and Smola 2002).

As studied and discussed by Mountrakis et al. (2011), SVMs have been extensively employed in remote sensing and geospatial data analysis due to their ability to use small training data sets, often resulting a higher classification accuracy than the traditional methods (Mantero et al. 2005). For instance, SVM has been used in road extraction from IKONOS imagery by (Huang and Zhang 2009) assessing the influence of the slope/aspect of the terrain on the forest classification accuracy (Huang et al. 2008), a crop classification task (Wilson et al. 2004), and many more factors.

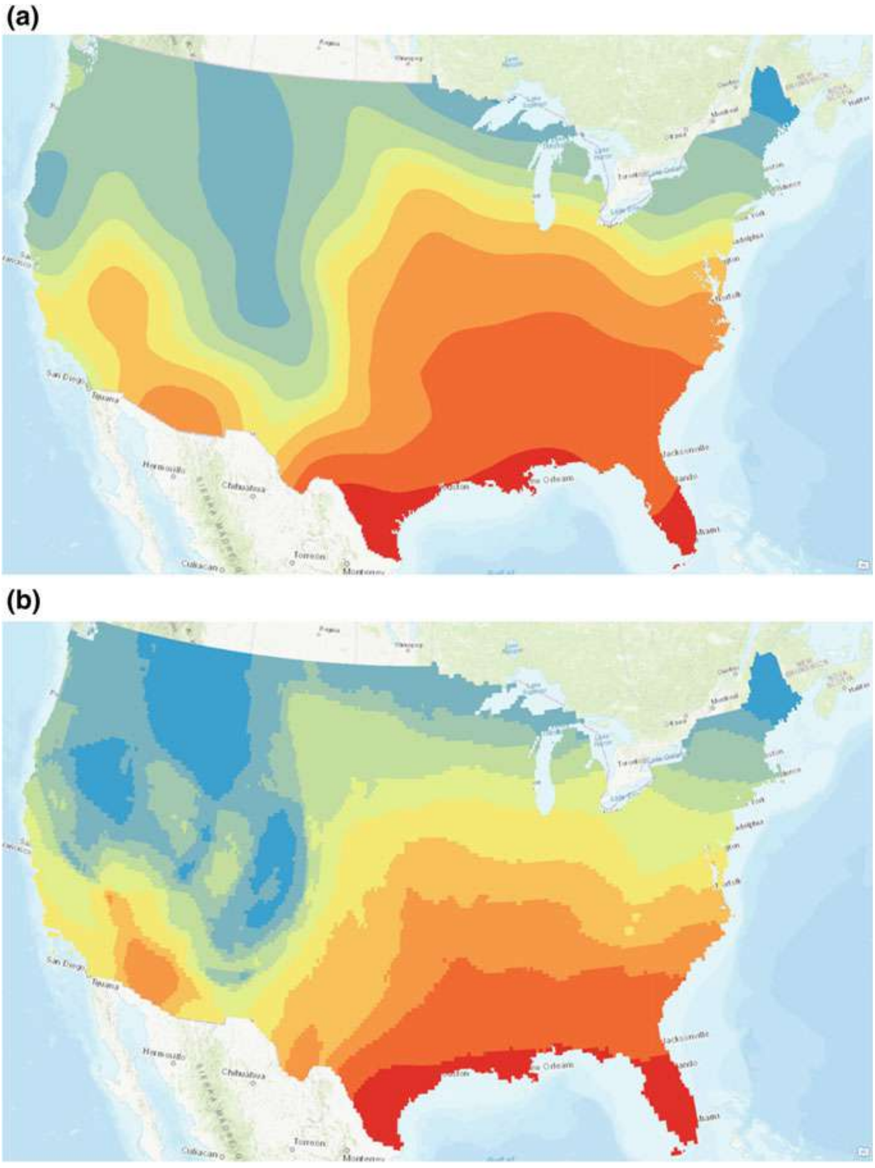


Fig. 10.4 **a** Downscaled temperature profile using GWR **b** downscaled temperature profile using a random forest regressor

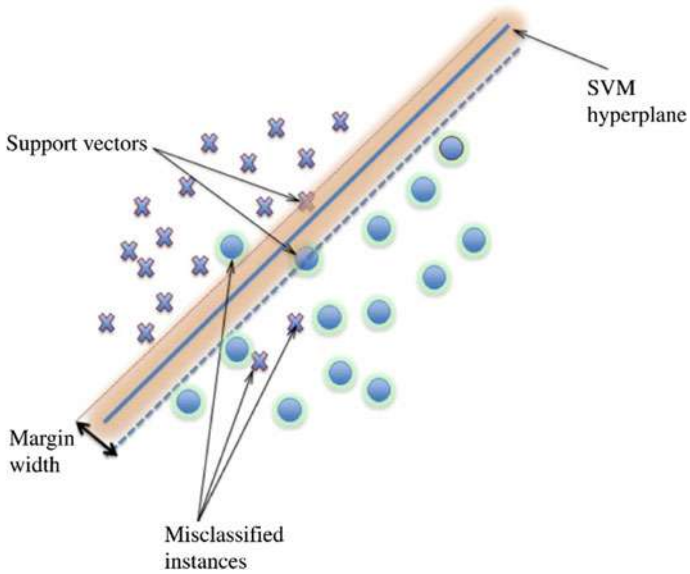


Fig. 10.5 SVM attempts to distinguish two categories of data by a hyperplane. Image from Mountrakis et al. (2011)

10.2.1.4 Active Contours and Active Shapes

Active contours or snakes have been developed with the aim of finding important features in an image by fitting a curve to the edges and lines of an image (Kass et al. 1988). Active contours are a set of energy-minimizing splines that are guided by external forces from the image. Snakes have been used extensively in geospatial image processing to detect features such as roads and buildings.

Active contours were later extended to active shapes to accommodate specific patterns in a set of objects and identify only those that are present in the training data (Cootes et al. 1995). In essence, they are very similar to active contours, but active shapes can only deform and fit the data that is consistent with the training set. Both active shapes and active contours have been extensively used in different applications of remote sensing and geoscience, such as object extraction (Liu et al. 2013), lane detection (Heij et al. 2004), and road extraction (see Fig. 10.6) (Kumar et al. 2017; Laptev 1997).

10.2.2 Unsupervised Learning

Unsupervised learning aims to infer the distribution of $P(X)$ in Eq. 10.1. Unlike supervised learning, $P(Y|X)$ or $P(X, Y)$ is not employed (Hastie et al. 2001). Thus,

Fig. 10.6 Active contours used to extract roads. Image taken form Laptev (1997)

unsupervised learning does not utilize any training dataset that contains information on $P(X, Y)$. One of the most common uses of unsupervised learning in geospatial analysis is in defining clusters and regions. These two terms differ, as clustering refers to defining groups based on value similarity in the data whereas regionalization performs clustering under spatial constraints (Duque et al. 2007). Both of these unsupervised learning approaches have wide applications (Duque et al. 2007; Hastie et al. 2001; Mitchell 1997; von Luxburg 2010). Most clustering and regionalization methods require definition of $\boldsymbol{k}$, the number of clusters to divide $\boldsymbol{X}$ into. There are extensive surveys of clustering and regionalization in the literature for readers to refer to (Duque et al. 2007; Jain et al. 1999).

10.2.2.1 SKATER Algorithm

As discussed in Chap. 8, the K-means algorithm (Macqueen 1967) aims to partition X into k groups and minimize the intergroup dissimilarity with the assumption that minimal intergroup dissimilarity corresponds to distinct groups. K-means seeks to create groups that consist of similar elements, ensuring that dissimilar elements are assigned to different groups. Mathematically:

$$\boldsymbol{\mu}(\boldsymbol{x}) = \boldsymbol{argmin}_{\mathrm{C}} \sum_{i=1}^{k} \sum_{x \in c_i} \left\| \boldsymbol{x} - \bar{\boldsymbol{C}}_{\iota} \right\|^2 \tag{10.4}$$

where $\boldsymbol{C} = \{\boldsymbol{C}_1, \boldsymbol{C}_2, \ldots, \boldsymbol{C}_k\}$ is the group of clusters, with a cluster $\boldsymbol{c}_m$ consisting of a subset of X and $\boldsymbol{c}_1 \cup \boldsymbol{c}_2 \cup \cdots \cup \boldsymbol{c}_k = \boldsymbol{X}$. K-means has various uses in geospatial analysis, including detecting patterns in traffic accidents (Anderson 2009), analyzing landslides (Keefer 2000) and creating labels by clustering topo-climatic data (Burrough et al. 2001).

The SKATER algorithm is a regionalization algorithm that imposes graph-based spatial constraints on the k-means algorithm (Assunção et al. 2006). Unlike Lloyd's algorithm (Lloyd 1982), SKATER only assigns spatially contiguous and similar objects to the same cluster. Regionalization has vast uses in geospatial analysis, including analysis of gerrymandering, healthcare services (Church and Barker 1998) and resource allocation (Or and Pierskalla 1979).

Clustering and regionalization were applied to the same dataset to juxtapose the types of patterns they expose in the data and the resulting understanding gained using these two methods. The average temperature in the United States in June 2012 was used. The resulting clusters and regions are displayed below.

The regionalization and clustering results in Fig. 10.7 show similarities in the overall temperature patterns, which change N-S in the eastern portion of the US and W-E in the western portion. Notably, the k-means result in Fig. 10.7b displays isolated patches whereas the regionalization result has spatially contiguous regions. Due to the constrained optimization scheme to satisfy the spatial constraints, the regions defined by regionalization have a higher variance than those in the k-means result. However, both maps display similarities in the temperature and the extent to which these similarities can be aggregated into homogeneous zones.

10.2.2.2 Autoencoders

Another very useful and common machine learning technique is autoencoders (Rumelhart et al. 1985). In an autoencoder, the data passes through a bottleneck, where the bottleneck is a lower representation of the same data. Autoencoders are made of two neural networks called the encoder and decoder (Fig. 10.8). The encoder receives data D, maps it to a lower space and obtains L; a decoder receives L, maps it back to the same dimension of D and obtains D'. The distance between D and D', which is called the reconstruction loss, should be minimized. A direct application of autoencoders is in compression, in which one can reduce the dimension of D to L and work with L and the decoder instead of the data D in its native resolution. Autoencoders have also been used in geospatial applications to find water bodies (Zhiyin et al. 2015) or denoise satellite images (Liang et al. 2017).

Machine learning techniques are not limited to the list of applications and methods provided here. Several variations of these methods as well as many other standalone techniques have been successfully employed in the Digital Earth, geoscience and remote sensing fields. For a more in-depth and comprehensive study, refer to the work of Lary et al. (2016).

10.2.3 Dimension Reduction

There have been extensive efforts to learn the patterns and forms that data sets contain. It is possible to predict the behavior of a data set and/or compress the data set into a more compact form for transmission, storage, and retrieval. In addition to autoencoders that can be used for dimensionality reduction, one of the easiest methods for compression and dimensionality reduction for a given data set and subsequent prediction of its behavior for unknown data points is *linear regression*.

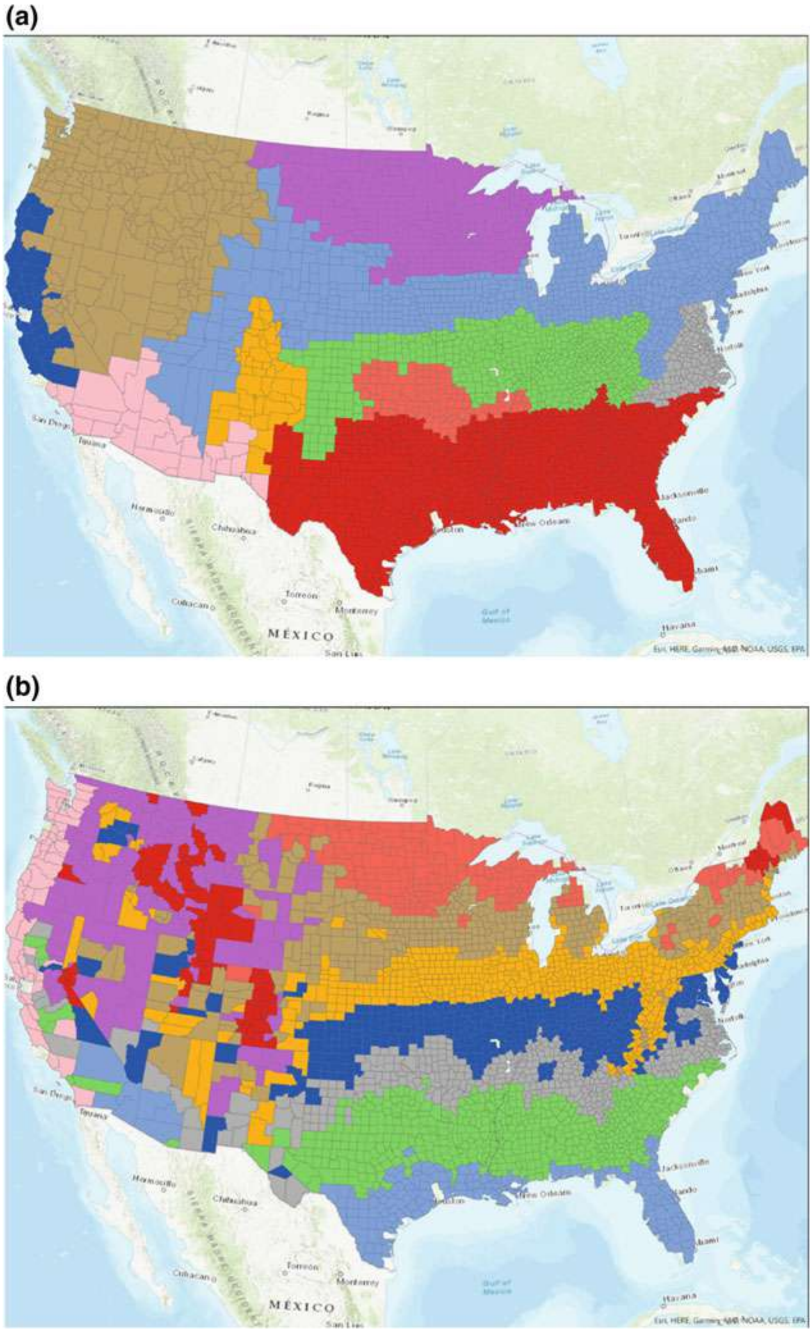


Fig. 10.7 **a** Temperature regions defined by SKATER **b** temperature regions defined by k-means

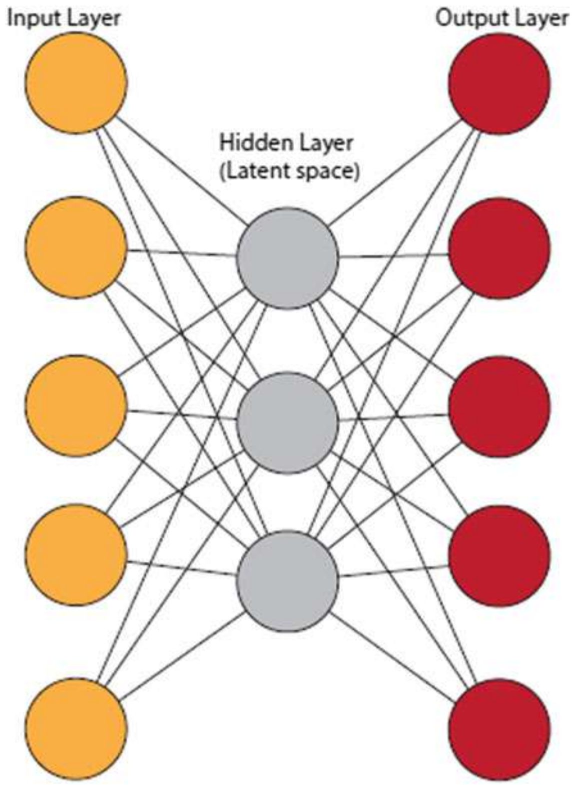


Fig. 10.8 The autoencoder passes the data (yellow neurons) through an encoder to learn a lower dimension (hidden/latent space; gray neurons) representation of the data. The decoder attempts to reconstruct the data (red neurons) as closely as possible to the given data

In the 2D case of this method, the data points attain two coordinates, and the line that best represents these data sets is considered as the model representative of the data. The best representation can have different meanings, including the line that has the smallest least square distance with all the data points. Regression, linear or nonlinear, has been a great tool to analyze spatial data. Belae et al. (2010) provided a survey of regression techniques used to represent and analyze spatial datasets. For Digital Earth platforms, Mahdavi-Amiri et al. (2018) combined regression with a wavelet to transmit quantitative datasets on a discrete global grid system (DGGS).

10.2.3.1 PCA

Another form of linear representation of a data set is principal component analysis (PCA). In this representation, the covariance matrix of the data is initially formed by applying the inner product of a data matrix A in its transpose $\left(Cov = A^T A\right)$. The eigenvectors of the covariance matrix, λ_i, represent the main trends of the data. If we have a data set forming an ellipsoid in 2D, the eigenvectors are the two main axes of the ellipsoid. Figure 10.9 represents PCA in 2D. PCA has been extensively used in

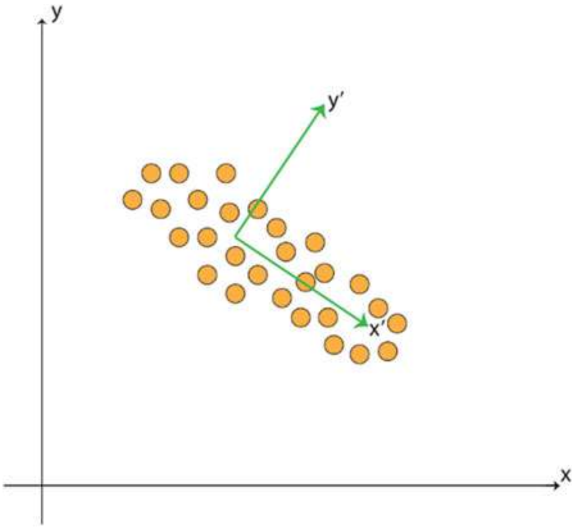


Fig. 10.9 PCA finds the main trends of the data. The data points illustrated in yellow have two main trends x′ and y′ that are the eigenvectors associated with the largest eigenvalues of the covariance of the data

many applications including computer graphics, computer vision, and data science. PCA has been used in different applications related to geospatial data representation and geospatial data analysis (Demšar et al. 2013). For instance, PCA has been successfully used to study drought areas (Gocic and Trajkovic 2014), evaluate water quality (Parinet et al. 2004), and distinguish vegetation (Panda et al. 2009).

10.2.3.2 SVD

Singular value decomposition (SVD) is a decomposition that reveals important information about a matrix. In SVD, a matrix A is decomposed into the form USV^T, in which U and V are two rotation matrices and S is a diagonal scale matrix with values called the singular values, σ_i, of matrix A. There is a direct connection between PCA and SVD because the singular values of the singular value decomposition of data matrix A are the square root of the eigenvalues of the covariance matrix that is found in PCA $\left(\sigma_i = \sqrt{\lambda_i}\right)$. To compress or denoise data, it is possible to zero out small eigenvalues obtained by SVD and keep important portions of the data. SVD has been extensively employed in image processing applications (Sadek 2012). It has also been used in geospatial applications. For instance, Wieland and Dalchow (2009) used SVD to detect landscape forms, and Dvorsky et al. (2009) used SVD to determine the similarity between maps.

10.2.3.3 Evolutionary and Agent-Based Techniques

Evolutionary and Agent-based techniques have also been extensively used to perform analyses of geospatial data sets. Two important algorithms are genetic algorithms (GAs) and ant colony optimization (ACO).

In GAs, a set of random solutions is initially produced and these solutions are considered parents to make a new generation of solutions based on three rules: **Selection** rules that select parents based on their fitness, **Crossover** rules that combine two parents to generate children for the next generation, and **Mutation** rules that apply random changes to parents to form children (Mitchel 1998). GAs have been used in many applications in geospatial data analysis such as road detection (Jeon et al. 2002) and satellite image segmentation (Mohanta and Binapani 2011).

ACO is an optimization technique that works based in an agent-based environment. In this stochastic environment, the ants are agents that walk over a certain solution path and leave a track called a pheromone. Paths with more pheromone are usually more optimal (shortest) than others, and they attract more agents. A classic problem that can be solved by ACO is the travelling salesman problem. ACO has been successfully employed to solve other types of hard problems including those involving geospatial data analysis. For instance, ACO has been used for path planning considering traffic (Hsiao et al. 2004) and road extraction from raster data sets (Maboudi et al. 2017).

10.3 Deep Learning

When a large amount of data is involved and/or a complex model for representing the data is used, it is common to employ deep learning methods (Goodfellow et al. 2016). Digital earth data represents a massive amount of data, for example, high-precision digital elevation models or aerial photography. Because the rules that produce this kind of data are very complex and involve many natural or human processes, it can be difficult to apply standard learning models or algorithms and retain this complexity. Thus, the deep models described in this section are relevant.

10.3.1 Convolutional Networks

Deep learning has been popularized by image processing applications. In this context, the processed data is arranged into a regular grid and is adapted to so-called convolutional layers. Data extracted from Digital Earth can be of this nature by construction. For example, raster data such as digital elevation models or aerial photography images are already arranged into regular grids and can be processed out of the box with convolutional layers. Convolutional neural networks rely on the fact that the same processing can be applied to different parts of the image. Traditional

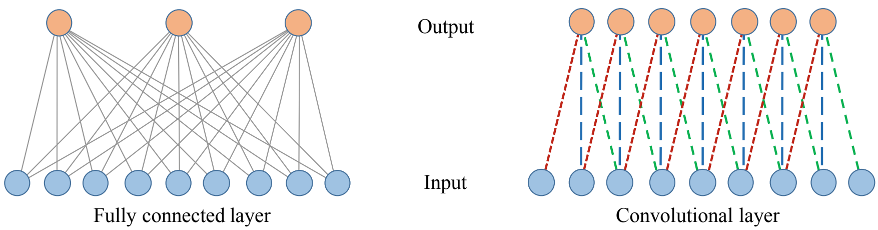


Fig. 10.10 Convolutional layers use fewer coefficients and are spatialized

fully connected schemes for neural network layers use many coefficients that can be spared with convolutional layers and used in other features. Figure 10.10 compares the principle of a convolutional layer to that of a traditionally fully connected layer. Both examples show an input of size 9. While a fully connected layer uses 27 coefficients to produce an output of size 3, the convolutional layer can produce 9 outputs from only 3 different coefficients. This means that the same feature extraction is performed but at different locations, which is relatively close to traditional convolution in the discrete domain.

Recently, a convolutional network was used to infer the super-resolution of a digital elevation model by using aerial photography (Argudo et al. 2018). Figure 10.11 shows the architecture of this network. This work comes from the observation that publicly available high-resolution DEMs (resolution less than 2 m) do not cover the full Earth whereas it is possible to find high-resolution imagery (orthophotos) with good coverage of the Earth. Many applications require a fine resolution for the DEM, and Argudo et al. proposed inserting details into a coarse DEM using inferred information drawn from the high-resolution orthophoto of the same footprint (Fig. 10.12). Basically, the method produces a DEM with 2 m precision from a DEM with 15 m precision and an orthophoto with 1 m precision. To produce this result, a fully convolutional network was used.

In the literature, a full system to automatically infer street addresses from satellite imagery was proposed (Demir et al. 2018a). One step that must be performed is the extraction of roads from the satellite images. This was done using a modified

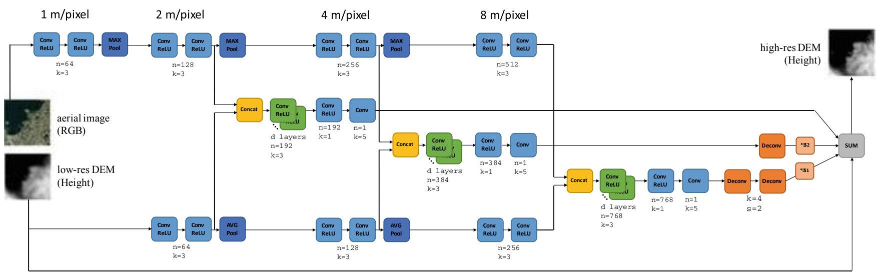


Fig. 10.11 A fully convolutional network was used to infer the high-resolution DEM from its coarse version and the high-resolution orthophoto (courtesy of O. Argudo et al.)

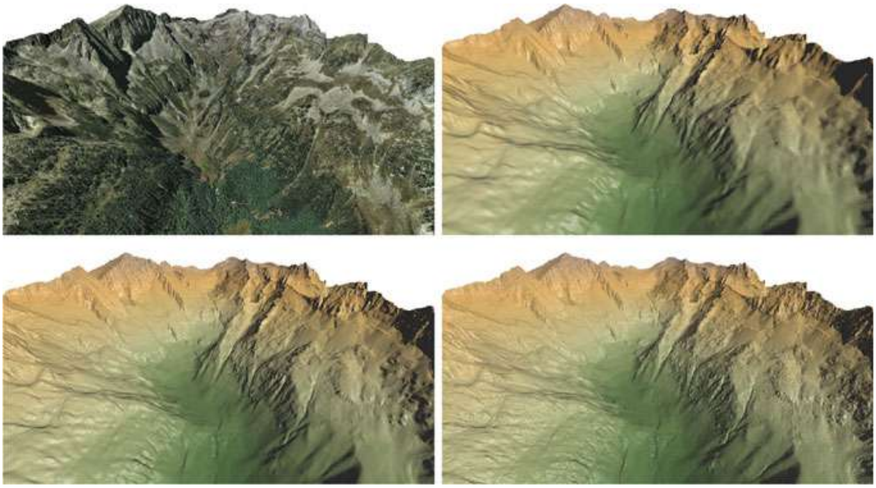


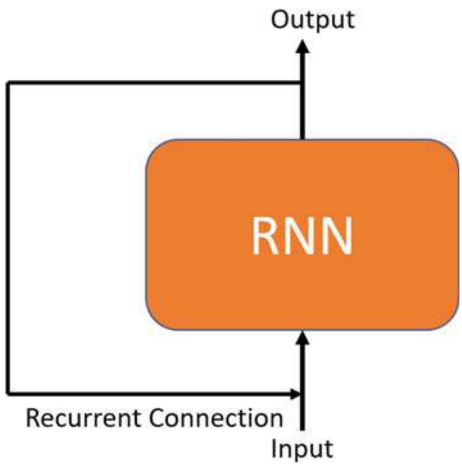
Fig. 10.12 Super-resolution of a 15 m precision DEM (top right) using an orthophoto (top left). Result (bottom left) and the ground truth reference (bottom right) (courtesy of O. Argudo et al.)

version of SegNet, a convolutional network primarily used for image segmentation. In this architecture, the input and output resolutions are identical, and the network consists of several encoder layers that decrease the resolution followed by decoders that increase the resolution. The network is trained using manually labeled 192 × 192 pixel images, in which a binary road mask is associated with each pixel of the image to indicate if the pixel belongs to a road or not. Figure 10.13 shows an example of the results obtained in automatic extraction of the road information compared with the ground truth.

More generally, automatic processing of satellite images with a deep learning approach appears to be very efficient in segmentation and feature extraction. The DeepGlobe project (http://deepglobe.org) aims to challenge authors to use deep learning for three applications: road extraction, building detection and land cover classification (Demir et al. 2018b).

Fig. 10.13 Automatic extraction of the road mask (right) from the satellite image (left), compared with the ground truth road network (center) (courtesy of I. Demir et al.)

Fig. 10.14 The schematic of a recurrent neural network



10.3.2 Recurrent Neural Networks

While convolutional neural networks and dense neural networks work well for static data in which there is no sense of time, a **r**ecurrent **n**eural **n**etwork (RNN) (Jain and Medsker 1999) processes data by iterating through the input elements and maintaining a state that contains information relative to what it has seen until then. An RNN is a neural network with an internal loop (see Fig. 10.14). The state of the RNN is updated between processing independent sequences; therefore, we still consider one data sequence as a single data point in the network. The difference is that this data point is not processed in a single step as opposed to those in dense or convolutional neural networks. In an RNN, the network internally loops over sequence elements until it learns the flow of the data. An RNN is helpful when dealing with a temporal data set. In geospatial data analysis, an RNN has been recently applied in interesting applications such as correction of satellite image classification (Maggiori et al. 2017) and land cover classification (Ienco et al. 2017). Since many types of geospatial data sets such as weather, satellite images, or seasonal animal behavior have timing attached to them, we expect that RNNs will be widely used in the analysis of geospatial data sets in the near future and that Digital Earth will benefit from such networks.

10.3.3 Variational Autoencoder

Deep neural networks are useful to analyze data sets and are also helpful in generating new data sets. It is possible to consider two deep neural networks as the encoder and decoder of an autoencoder and produce a latent space that represents the data. Using only L and an encoder, we can reproduce a lossy representation of D. However, it is not possible to pick a vector in L and expect to reproduce a meaningful result by feeding it to the encoder because the distribution of L is unknown if autoencoders are used. In **v**ariational **a**uto**e**ncoders (VAEs) (Kingma and Welling 2014), in addition to the compression loss, another loss is minimized that forces the L to be a Gaussian distribution. Thus, VAEs can be used as a generative neural network in which one can sample the Gaussian distribution and feed it to the encoder to generate a new shape that does not necessarily belong to the training data set. Although VAEs have potential to generate data and learn low-dimensional data for geospatial data sets, VAEs have not been extensively tested for geospatial data analysis and generation.

10.3.4 Generative Adversarial Networks (GANs)

Similar to VAEs, **g**enerative **a**dversarial **n**etworks (GANs) (Goodfellow et al. 2014) are also generative models. GANs consist of a pair of networks that have two different and adversarial roles. These networks have a convolutional architecture and are often complex to retain the complexity of the underlying models. The first network is a generator that we denote as G, which attempts to generate the best result, for example, an image. Then, the second network takes the image as input and tries to infer if it is a generated image or not. This second network is called a discriminator and we denote it as D. Both G and D are trained alternatively. The objective of G is to fool D whereas D aims to avoid being fooled by G. The strength of this kind of adversarial formalism is that it is equivalent to use of a very complex function to train the generator G (encoded into the discriminator), far more complex than traditional distance would be.

Conditional **G**ANs (cGANs) are GANs with a particular setup in which the discriminator is trained to recognize the matching between an input image A and an output B whereas a traditional GAN only tests the plausibility of the output without any knowledge of the input. The training principle of a cGAN is explained in Fig. 10.15.

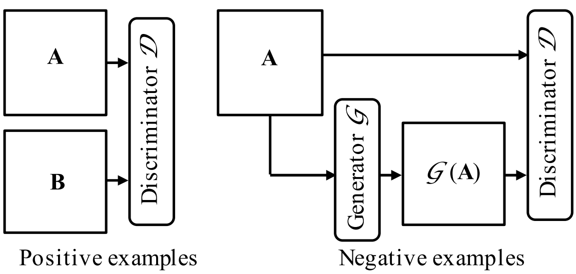


Fig. 10.15 cGAN principle: a training pair (A, B) is used to learn positive examples. For negative examples, only A is used together with the generator to form the pair (A, G(A))

Conditional GANs have recently been used to automatically generate digital elevation models from user sketches (Guérin et al. 2017). The user sketches the river network, the crests and some altitude cues and obtains a plausible terrain that matches the given constraints, based on a training dataset made of sketch/terrain pairs. The method consists of building such a dataset by extracting the sketch from a real-world terrain. The difficulty of this kind of setup is to automatically build a sketch that is compatible with user sketches, i.e., similar to what a user would draw. Building a sketch that is too close to the terrain features will force the user to draw very precisely, which is not relevant in a sketching context but would be useful in a reconstruction process. The digital elevation model must be simplified to produce simpler features. In their work, Guérin et al. propose initially downsampling the digital elevation model and then smoothing it. This coarse digital elevation model is then processed by a flow simulation, from which the skeleton is extracted. The same process is applied to extract ridges. This feature extraction is illustrated in Fig. 10.16.

The training dataset is formed of pairs that describe the matching between the sketch and the terrain. Figure 10.17 gives examples of such pairs. To create a more pliable terrain synthesizer, the sketches randomly include one, two or the three features among the river lines, crest lines and altitude cues.

Figure 10.18 shows examples of outputs produced by the DEM generator from sketches. The results were obtained by using training from a DEM extracted from the NASA SRTM dataset at 1 arc-second from different locations in the United

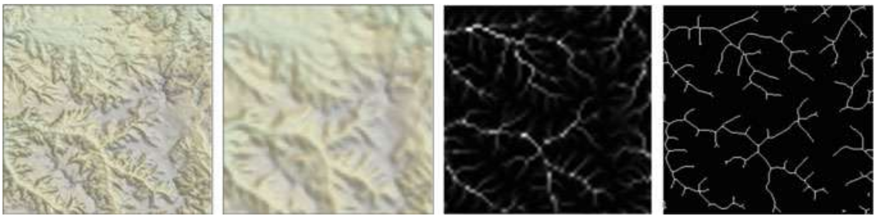


Fig. 10.16 Training database examples

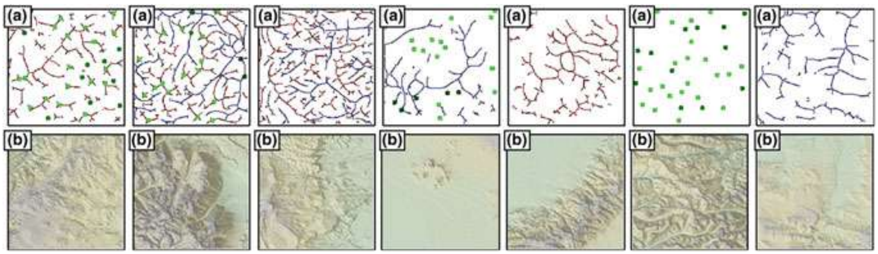


Fig. 10.17 Training database examples. Training pairs are formed by a sketch (**a**) and an associated DEM (**b**). Sketches can feature river lines (blue), crests (red) and altitude cues (green)

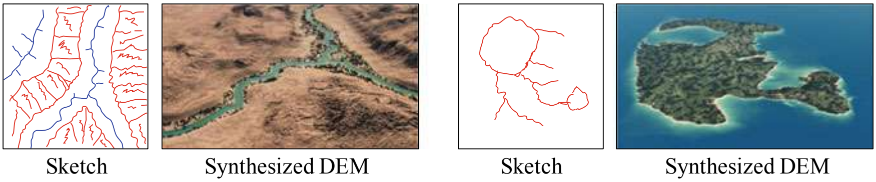


Fig. 10.18 Examples of generated digital elevation models from simple sketches. A canyon generated using river and crest lines (left). A volcanic island generated using only crest lines (right)

States. In the same article, the authors proposed the use of the same principle to automatically generate digital elevation models from a single level set sketch. They also described examples of automatic void filling in digital elevation models. Finally, because cGANs can embed very complex models, they used it to mimic an erosion process.

10.3.5 Dictionary-Based Approaches

Approaches based on base function decompositions have intrinsic limitations. Base functions are usually used because they have orthogonality properties that lead to an efficient decomposition. Selecting the base can be difficult because it heavily depends on the nature of the signal. Thus, it can be a viable option to use dictionary-based descriptions. A signal is represented as a linear combination of atoms from a dictionary. Atoms do not need to have special properties such as orthogonality. They are typically chosen directly from the data by picking the most representative signals or by using an optimization. A survey of dictionary-based methods for 3D modeling was conducted by Lescoat et al. (2018). One of the applications of dictionary-based modeling is called sparse modeling, which adds an additional constraint on the number of atoms used to represent the final signal, called *sparsity*.

10.3.5.1 Dictionary Decomposition

Given a dictionary, the decomposition of a signal consists of finding the best atom, i.e., the atom that maximizes the projection. Then, the same process is applied iteratively to the residual until reaching the target sparsity. This process is called matching pursuit and was introduced by Mallat and Zhang (1993). This decomposition algorithm was further improved by Cai and Wang (2011) by introducing the *Orthogonal Matching Pursuit* (OMP) algorithm. The main difference is that the best decomposition of the already-found atoms is recomputed after each new atom is found.

Fig. 10.19 An example of terrain amplification that adds plausible details from a given exemplar using a dictionary-based approach. The original terrain had a precision of 1 km, and successive amplifications by a factor of 4 increase the precision to 4 m

10.3.5.2 Dictionary Optimization

One aim of dictionary based approaches is to find a dictionary that is adapted to a given context or set of signals. This can be done by an optimization process. One goal of this optimization is to minimize the reconstruction error, for example, by computing an L^2 distance between the reconstructed signal and the original. It is common to add a constraint on the type of decomposition, for example, by setting a maximum sparsity. Unfortunately, the optimization problem under this type of constraint is too difficult to solve in an optimal way. Heuristics have been proposed that lead to good results with a relatively low cost. K-SVD is one of these algorithms (Aharon et al. 2006), which consists of iterating between two steps. The first step consists of optimizing the decomposition, which can be done using a standard OMP algorithm. The second step optimizes the dictionary with respect to the previously computed decomposition. The two steps are repeated until a number of iterations is reached or a given error is obtained.

Several applications of sparse modeling with terrains have been proposed by Guérin et al. (2016) and Argudo et al. (2018). The terrain is decomposed into patches that compose input signals. A so-called amplification process is used to introduce plausible details into the terrain by mapping between low-resolution and hi-resolution atoms. The dictionary is drawn from an exemplar terrain at high resolution and automatically transformed into low resolution by a trivial downsampling process. The amplification algorithm simply decomposes the patches from a given terrain in the low-resolution dictionary and uses the corresponding high-resolution atoms to reconstruct it. Because the dictionary has been extracted from real terrain, the added details are plausible and realistic, as shown in Fig. 10.19.

10.3.6 Reinforcement Learning

Reinforcement **l**earning (RL) is a powerful learning method in dynamic environments (Sutton and Barto 1998). In RL, there is usually an agent in an environment and the agent receives rewards based on its actions. The final goal is to learn how to take actions to maximize the rewards. At any time t, an environment is defined by states S_t in which an agent can take action A_t and change the environment state to S_{t+1}. When

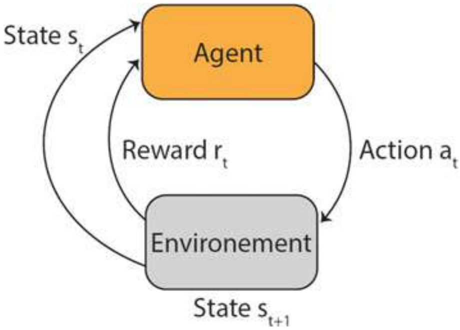


Fig. 10.20 An agent receives state s_t, performs an action and receives reward r_t from the environment. The state of the environment changes to s_(t+1). This process continues until a terminal state is achieved

the agent takes action A_t, the environment receives a reward r_t. These iterations continue until the environment reaches a terminal state (Fig. 10.20). Examples of applications that RL can be extremely useful for are games or robot locomotion in which more points and more stable states are the rewards of the game and locomotion environments, respectively.

RL has also been used in applications in GIS and geospatial data analysis. For instance, RL has been used to model land cover changes (Bone and Dragicevic 2009). With recent advances in RL and the growth of computational power, we expect that RL will receive more attention from the GIS and Digital Earth communities. For instance, one application of RL can be to simulate the behavior of endangered species in different simulated environments.

10.4 Discussion

In the past, machine learning has seen hypes and winter seasons. It started with symbolic AI in the 1960s, which claimed the ability to make machines with intelligence comparable to an average human being in less than a decade. However, people soon realized that they were far from reaching that point. In the 1980s, with the rise of *expert systems*, similar hype was seen in the area of machine learning, followed by a winter season due to the lack of generality of expert systems and their high maintenance costs (Chollet 2017). Recently, deep learning methods became popular again and showed great success in different areas of computer science including geospatial analysis, which is an important portion of Digital Earth platforms. Deep learning will likely continue to grow and be applied more in this field, especially because of the availability of computational power and big data sets that help create more powerful models. However, deep learning cannot solve all problems. For instance, current deep learning models are unable to solve problems that require reasoning or long-term planning (Chollet 2017). Deep learning models work extremely well in mapping an input to a desired output with very little human-level knowledge about the input or output and their effect on industry and science will probably remain for a very long time. There is plenty of discussion about the future of deep learning and AI, notably by its great pioneers such as Lecun et al. (2015) and in the European perspective on AI (Craglia et al. 2018).

Artificial intelligence and particularly machine learning and deep learning have great potential to contribute to the generation, analysis, and management of geospatial data sets. Digital Earth should benefit from such opportunities, as a place holder to represent such data sets and a platform to analyze them. Since Digital Earth is constantly receiving geospatial data sets, a successful Digital Earth should use reliable, fast, and comprehensive techniques to manage and make use of such data. Deep Learning techniques show promise in these directions. However, there are still issues in their use in Digital Earth platforms that must be addressed. In the following sections, we discuss some of these issues.

10.4.1 Reproducibility

If a technique such as a deep neural network produces particular results, such results should be reproduceable by others. Placing the code on GitHub and providing free access to data sets have been helpful for this issue. However, there are still some issues, especially when the data are owned by a company or the network was designed by an industrial team. In particular neural network architectures, randomness can be included, usually to improve the training. When this randomness is also present in the operational network, it can disrupt the reproducibility of results.

10.4.2 Ownership and Fairness

Ownership of artifacts provided by machine learning techniques is also heavily under question. If a person with almost no knowledge about a network takes information from available sources, modifies a few parameters, takes data from an available source and produces something unique or obtains a certain analysis, who is the owner of such results? The data owner, developer of the network, or the person who combined these ingredients? In more serious scenarios, who is at fault when a system that works based on machine learning techniques makes a catastrophic mistake or performs a discriminatory action that may involve racism or sexism? Another question is whether data sets and computation power are available to everyone, i.e., do we have "data democratization"? Fortunately, the wealth of free access data sets and code bases along with cheap computational power such as Amazon Web Services (AWS) have resolved some of these issues but we are still far from perfect.

10.4.3 Accountability

Due to the nature of some algorithms involved in machine learning, it usually cannot be used in contexts where accountability is a strong constraint. This is especially the case with deep neural networks where a lot of information is hidden in the layers, which can lead to unexpected and unwanted results. Conversely, traditional machine learning methods such as linear regressions or PCA are very reliable even if they are limited in terms of applications. Reasonably, one could consider using deep learning methods only when traditional methods fail or are lacking.

10.5 Conclusion

In conclusion, we provided a sampling of artificial intelligence techniques and their applications in geospatial data generation, analysis, and management. We discussed how AI can be beneficial for generating new terrain data sets, identifying roads and analyzing various geospatial data sets such as satellite imagery. AI techniques and deep learning methods appear very promising. Extensive research on these topics will likely make them even more suitable for use in different domains including geospatial analysis and Digital Earth. However, these techniques are unfortunately standalone and have not been integrated into a Digital Earth platform that makes use of such techniques. Appropriate artificial intelligence techniques should be meticulously included in Digital Earth, considering their pros and cons including fairness and bias to provide interactive, comprehensive and meaningful analysis to users.

References

Aharon M, Elad M, Bruckstein A (2006) K-SVD: an algorithm for designing overcomplete dictionaries for sparse representation. IEEE Trans Signal Process 54(11):4311–4322
Anderson TK (2009) Kernel density estimation and K-means clustering to profile road accident hotspots. Accid Anal Prev 41(3):359–364
Argudo O, Chica A, Andujar C (2018) Terrain super-resolution through aerial imagery and fully convolutional networks. Comput Gr Forum 37(2):101–110
Assunção RM, Neves MC, Câmara G et al (2006) Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. Int J Geogr Inf Sci 20(7):797–811
Beale CM, Lennon JJ, Yearsley JM et al (2010) Regression analysis of spatial data. Ecol Lett 13(2):246–264
Bone C, Dragićević S (2009) Defining transition rules with reinforcement learning for modeling land cover change. Simulation 85(5):291–305
Breiman L (1996) Bagging predictors. Mach Learn 24(2):123–140
Breiman L (2001) Random forests. Mach Learn 45(1):5–32
Bunge W (1966) Gerrymandering, geography, and grouping. Geogr Rev 56(2):256–263

Burrough PA, Wilson JP, van Gaans PFM et al (2001) Fuzzy k-means classification of topo-climatic data as an aid to forest mapping in the Greater Yellowstone Area, USA. Landsc Ecol 16(6):523–546
Cai TT, Wang L (2011) Orthogonal matching pursuit for sparse signal recovery with noise. IEEE Trans Inf Theory 57(7):4680–4688
Chollet F (2017) Deep learning with python. Manning Publications Company, Shelter Island
Church J, Barker P (1998) Regionalization of health services in Canada: a critical perspective. Int J Health Serv 28(3):467–486
Cootes TF, Taylor CJ, Cooper DH et al (1995) Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59
Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297
Cortés U, Sangüesa MS-MR, Comas J et al (2001) Knowledge management in environmental decision support systems. AI Commun 14(1):3–12
Craglia M, Annoni A, Benczur P et al (2018) Artificial intelligence: a european perspective. Publications Office of the European Union, Luxenbourg
Cutler DR, Edwards Jr. TC, Beard KH et al (2007) Random forests for classification in ecology. Ecology 88(11):2783–2792
Demir İ, Hughes F, Raj A et al (2018a) Generative street addresses from satellite imagery. ISPRS Int J Geo-Inf 7(3):84
Demir I, Koperski K, Lindenbaum D et al (2018b) DeepGlobe 2018: a challenge to parse the Earth through satellite images. Comput Vis Pattern Recognit. https://arxiv.org/abs/1805.06561 Accessed 2 Aug 2019.
Demšar U, Harris P, Brunsdon C et al (2013) Principal component analysis on spatial data: an overview. Ann Assoc Am Geogr 103(1):106–128
Duque JC, Ramos R, Suriñach J (2007) Supervised regionalization methods: a survey. Int Reg Sci Rev 30(3):195–220
Dvorský J, Snášel V, Voženílek V (2009) Map similarity testing using matrix decomposition. In: 2009 international conference on intelligent networking and collaborative systems, IEEE, Barcelona, 4–6 November 2009
Egenhofer MJ, Mark DM (1995) Naive geography. In: Frank AU, Kuhn W (eds) Spatial information theory A theoretical basis for GIS. Springer, Heidelberg, pp 1–15
Fonseca FT, Egenhofer MJ, Agouris P et al (2002) Using ontologies for integrated geographic information systems. Trans GIS 6(3):231–257
Fotheringham AS, Brunsdon C, Charlton M (2003) Geographically weighted regression: the analysis of spatially varying relationships. Wiley, Chichester
Gislason PO, Benediktsson JA, Sveinsson JR (2006) Random forests for land cover classification. Pattern Recognit Lett 27(4):294–300
Gocic M, Trajkovic S (2014) Spatiotemporal characteristics of drought in Serbia. J Hydrol 510:110–123
Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press, Cambridge
Goodfellow I, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C et al (eds) Advances in neural information processing systems 27 (NIPS 2014). NIPS, Montreal, pp 2672–2680
Gruber TR (1993) Toward principles for the design of ontologies used for knowledge sharing? Int J Hum-Comput Stud 43(5):907–928
Guérin E, Digne J, Galin E et al (2016) Sparse representation of terrains for procedural modeling. Comput Gr Forum 35(2):177–187
Guérin É, Digne J, Galin É et al (2017) Interactive example-based terrain authoring with conditional generative adversarial networks. ACM Trans Gr 36(6):1–13
Hastie T, Tibshirani R, Friedman JH (2001) The elements of statistical learning: data mining, inference, and prediction. Springer, New York, NY
Heij R, Helgers M, Kockelkorn W et al (2004) Snakes for lane detection. Image Vis Comput 269–280

Hsiao Y-T, Chuang C-L, Chien C-C (2004) Ant colony optimization for best path planning. In: IEEE international symposium on communications and information technology, 2004. ISCIT 2004, IEEE, Sapporo, 26–29 October 2004

Huang H, Gong P, Clinton N et al (2008) Reduction of atmospheric and topographic effect on Landsat TM data for forest classification. Int J Remote Sens 29(19):5623–5642

Huang X, Zhang L (2009) Road centreline extraction from high-resolution imagery based on multiscale structural features and support vector machines. Int J Remote Sens 30(8):1977–1987

Ienco D, Gaetano R, Dupaquier C et al (2017) Land cover classification via multitemporal spatial data by deep recurrent neural networks. IEEE Geosci Remote Sens Lett 14(10):1685–1689

Jain AK, Murty MN, Flynn PJ (1999) Data clustering: a review. ACM Comput Surv 31(3):264–323

Jeon B-K, Jang J-H, Hong K-S (2002) Road detection in spaceborne SAR images using a genetic algorithm. IEEE Trans Geosci Remote Sens 40(1):22–29

Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331

Keefer DK (2000) Statistical analysis of an earthquake-induced landslide distribution — the 1989 Loma Prieta, California event. Eng Geol 58(3):231–249

Kingma D, Welling M (2014) Auto-encoding variational bayes. Mach Learn. arXiv preprint arXiv:1312.6114

Kuipers BJ (1996) An ontological hierarchy for spatial knowledge. In: Proceedings of 10th international workshop on qualitative reasoning about physical systems. AAAI Press, Fallen Leaf, pp 113–120

Kumar P, Lewis P, McCarthy T (2017) The potential of active contour models in extracting roads from mobile laser scanning data. Infrastructures 2(3):1–16

Laptev I (1997) Road extraction based on snakes and sophisticated line extraction. Master's thesis, Royal Institute of Technology

Lary DJ, Alavi AH, Gandomi AH et al (2016) Machine learning in geosciences and remote sensing. Geosci Front 7(1):3–10

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521(7553):436–444

Lenat DB (1995) CYC: a large-scale investment in knowledge infrastructure. Commun ACM 38(11):33–38

Lescoat T, Ovsjanikov M, Memari P et al (2018) A survey on data-driven dictionary-based methods for 3D modeling. Comput Gr Forum 37(2):577–601

Liang P, Shi W, Zhang X (2017) Remote sensing image classification based on stacked denoising autoencoder. Remote Sens 10(1):16

Liu G, Sun X, Fu K et al (2013) Interactive geospatial object extraction in high resolution remote sensing images using shape-based global minimization active contour model. Pattern Recognit Lett 34(10):1186–1195

Lloyd S (1982) Least squares quantization in PCM. IEEE Trans Inf Theory 28(2):129–137

Maboudi M, Amini J, Hahn M et al (2017) Object-based road extraction from satellite images using ant colony optimization. Int J Remote Sens 38(1):179–198

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: le Cam LM, Neyman J (eds) Proceedings of the fifth berkeley symposium on mathematical statistics and probability, volume 1: statistics. University of California Press, Berkeley, pp 281–297

Maggiori E, Charpiat G, Tarabalka Y et al (2017) Recurrent neural networks to correct satellite image classification maps. IEEE Trans Geosci Remote Sens 55(9):4962–4971

Mahdavi-Amiri A, Alderson T, Samavati F (2015) A survey of digital earth. Comput Gr 53:95–117

Mahdavi-Amiri A, Alderson T, Samavati F (2018) Geospatial data organization methods with emphasis on aperture-3 hexagonal discrete global grid systems. Cartographica 54(1):30–50

Mallat SG, Zhang Z (1993) Matching pursuits with time-frequency dictionaries. IEEE Trans Signal Process 41(12):3397–3415

Mantero P, Moser G, Serpico SB (2005) Partially supervised classification of remote sensing images through SVM-based probability density estimation. IEEE Trans Geosci Remote Sens 43(3):559–570

McCarthy J (1988) Mathematical logic in artificial intelligence. Daedalus 117(1):297–311
Medsker L, Jain LC (1999) Recurrent neural networks: design and applications. CRC Press, Boca Raton
Mitchell T (1997) Machine learning. McGraw Hill, Burr Ridge
Mitchell M (1998) An introduction to genetic algorithms. MIT Press, Cambridge
Mohanta RK, Binapani S (2011) A review of genetic algorithm application for image segmentation Int J Comput Technol Appl 3(2):720–723
Mountrakis G, Im J, Ogole C (2011) Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 66(3):247–259
Nelder JA, Wedderburn RWM (1972) Generalized linear models. J Royal Statist Soc Ser A (Gen) 135(3):370–384
Or I, Pierskalla WP (1979) A transportation location-allocation model for regional blood banking. AIIE Trans 11(2):86–95
Panda SS, Hoogenboom G, Paz J (2009) Distinguishing blueberry bushes from mixed vegetation land use using high resolution satellite imagery and geospatial techniques. Comput Electron Agric 67(1):51–58
Parinet B, Lhote A, Legube B (2004) Principal component analysis: an appropriate tool for water quality evaluation and management—Application to a tropical lake system. Ecol Model 178(3):295–311
Rizzoli AE, Young WJ (1997) Delivering environmental decision support systems: software tools and techniques. Environ Model Softw 12(2):237–249
Rumelhart DE, Hinton GE, Williams RJ (1985) Learning internal representations by error propagation. Institute for Cognitive Science, University of California, San Diego, La Jolla
Rykiel EJ (1989) Artificial intelligence and expert systems in ecology and natural resource management. Ecol Model 46(1):3–8
Sadek R (2012) SVD based image processing applications: state of the art, contributions and research challenges. Int J Adv Comput Sci Appl 3(7):26–34
Schölkopf B, Smola AJ (2002) Learning with kernels: support vector machines, regularization, optimization, and beyond. MIT Press, Cambridge
Sowa JF (2000) Knowledge representation: logical, philosophical, and computational foundations. Brooks/Cole, Pacific Grove
Sutton R, Barto A (1998) Introduction to reinforcement learning. MIT Press, Cambridge
Vapnik V (1979) Estimation of dependences based on empirical data. Springer, New York
von Luxburg U (2010) Clustering stability: an overview. Found Trends Mach Learn 2(3):235–274
Wieland R, Dalchow C (2009) Detecting landscape forms using Fourier transformation and singular value decomposition (SVD). Comput Geosci 35(7):1409–1414
Wilby RL, Wigley TML (1997) Downscaling general circulation model output: a review of methods and limitations. Prog Phys Geogr: Earth Environ 21(4):530–548
Wilson MD, Ustin SL, Rocke DM (2004) Classification of contamination in salt marsh plants using hyperspectral reflectance. IEEE Trans Geosci Remote Sens 42(5):1088–1095
Zhiyin W, Long Y, Shengwei T et al (2015) Water body extraction method based on stacked autoencoder. J Comput Appl 35(9):2706–2709

Eric Guérin is Assistant Professor at INSA de Lyon, France. He received an engineering degree from INSA de Lyon in 1998 and a Ph.D. in Computer Science from Université Claude Bernard Lyon 1 in 2002. His research interests include procedural modeling of virtual worlds, natural phenomena and machine learning. He is head of the LIRIS/Geomod team since 2015.

Orhun Aydin is a researcher in Environmental Systems Research Institute and assistant professor in Spatial Sciences Institute at U.S.C. Prior to joining ESRI and USC, he obtained a Ph.D. and M.Sc. in Geostatistics from Stanford University. Orhun's research pertains to modelling, exploring and forecasting spatial and spatio-temporal phenomena and their applications for Earth Systems.

Ali Mahdavi-Amiri is a University Research Associate at Computing Science department in Simon Fraser University. His research interests are in Computer Graphics with specialty in Geometric Modeling, Geospatial Modeling and Visualization, Digital Fabrication, and Machine Learning. Ali has received several awards including NSERC Postdoctoral Fellowship, Graduate Achievement Award, Michael A. J. Sweeney Award, and Teaching Excellence Award.

Chapter 11
Internet of Things

Carlos Granell, Andreas Kamilaris, Alexander Kotsev, Frank O. Ostermann and Sergio Trilles

Abstract Digital Earth was born with the aim of replicating the real world within the digital world. Many efforts have been made to observe and sense the Earth, both from space (remote sensing) and by using in situ sensors. Focusing on the latter, advances in Digital Earth have established vital bridges to exploit these sensors and their networks by taking location as a key element. The current era of connectivity envisions that everything is connected to everything. The concept of the Internet of Things (IoT) emerged as a holistic proposal to enable an ecosystem of varied, heterogeneous networked objects and devices to speak to and interact with each other. To make the IoT ecosystem a reality, it is necessary to understand the electronic components, communication protocols, real-time analysis techniques, and the location of the objects and devices. The IoT ecosystem and the Digital Earth (DE) jointly form interrelated infrastructures for addressing today's pressing issues and complex challenges. In this chapter, we explore the synergies and frictions in establishing an efficient and permanent collaboration between the two infrastructures, in order to adequately address multidisciplinary and increasingly complex real-world problems. Although there are still some pending issues, the identified synergies generate optimism for a true collaboration between the Internet of Things and the Digital Earth.

Keywords Internet of Things · Geospatial standards · Smart scenarios

The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

C. Granell · S. Trilles (✉)
Institute of New Imaging Technologies, Universitat Jaume I, Castelló de la Plana, Spain
e-mail: strilles@uji.es

A. Kamilaris
Department of Computer Science, University of Twente, Enschede, The Netherlands

A. Kotsev
European Commission, Joint Research Centre, Ispra, Italy

F. O. Ostermann
Faculty of Geographic Information Science and Earth Observation (ITC), Department of Geo-Information Processing, University of Twente, Enschede, The Netherlands

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_11

11.1 Introduction

According to Jayavardhana (Gubbi et al. 2013), the term Internet of Things (IoT) was first coined by Kevin Ashton in 1999 in the context of supply chain management. Empowered by the latest advances in Information and Communication Technology (ICT), the IoT is revolutionizing the world, opening new possibilities and offering solutions that were unthinkable even only a few years ago. The concept of the IoT is highly multidisciplinary because it brings together a wide variety of technologies, protocols, applications, scenarios, and disciplines (Atzori et al. 2010; Gubbi et al. 2013). The International Telecommunication Union (ITU) Standardisation Sector defines it as '*a global **infrastructure** for the information society, enabling advanced services by interconnecting* (*physical and virtual*) *Things based on existing and evolving interoperable information and communication technologies*' (International Telecommunication Union 2018). As an infrastructure, the IoT can be seen as a broader system involving data, resources, standards and communication protocols as well as theoretical studies.

The pace of IoT development seems quite fast, with continuous proposals of new approaches, applications, and use case scenarios, increasing the presence of IoT in multiple and varied applications, and aspects of daily life. To date, smart devices constitute the IoT's most visible form, applied in a wide range of scenarios and sectors such as cities, industry, commerce, agriculture, home, and mobility. Although we are far from the 200 trillion smart devices as predicted by 2020 (Intel, n.d.), significant progress has been made in this direction. Estimates suggest that there will be 26 smart devices per person in 2020, 40.2% of which will be located in the business environment (termed Industry 4.0).

According to the Forbes analyst Daniel Newman (Newman 2017), the IoT is one of the most rapidly evolving trends today, especially in three development lines: the analytics arena, the development of edge computing, and the deployment of 5G networks. As 5G technology is progressively implemented and deployed (Shafi et al. 2017), the current analysis platforms will need adaptation in order to analyze effectively the large amount of data flows acquired, produced by IoT devices with increasingly more powerful built-in sensors and emerging real-time analysis functions, empowered even more by the rapid emergence and (parallel) development of edge computing (Shi et al. 2016).

Edge computing is a recent paradigm motivated by bandwidth limitations between the producer (smart objects) and consumer parts (cloud server), as well as the need for improved performance in computing and consumer smart objects. The main feature of edge computing is that data can be processed locally in smart devices rather than being sent to the cloud for further processing.

Like the IoT, Digital Earth (DE) also entails an **infrastructure**. Al Gore, at his famous speech in 1998 (Gore 1998), introduced the concept of a DE with the vision of extending the real Earth with a digital/virtual replica or counterpart. Over the last two decades, many geographic phenomena and observations have been converted to digital data to be used, analyzed, and visualized using digital tools such as virtual globes (Butler 2006). In this chapter, we use the term DE to refer to a network infrastructure that allows for the discovery, access, analysis, and processing of spatially referenced data. For more details on DE, we refer the reader to Schade et al. (2013). In particular, Schade et al. describe the origins and evolving concepts of terms such as DE, Geographic Information Infrastructures and Spatial Data Infrastructures, together with their theoretical and technical features.

This chapter takes a technological perspective focusing on the description of the current relationships between DE and the IoT, identifying ongoing efforts, potential synergies and bridges, as well as existing limitations and barriers that prevent both infrastructures from collaborating and communicating in practical terms. Instead of operating in parallel, scientists and researchers need the IoT and DE to work jointly by establishing an efficient and permanent collaboration to adequately address the multi-disciplinary nature and growing complexity of the pressing problems that characterize modern science.

The rest of the chapter is divided into five sections. In Sect. 11.2, we provide an overview of the most frequent definitions of the IoT, describe our working definitions throughout this chapter, and briefly review related work in the interplay of the IoT and the DE. In Sect. 11.3, we analyze the existing interplay between both infrastructures in the context of the main, high-level functions of DE. Then, an overview of relevant case studies across several smart scenarios in which the symbiosis of the IoT and DE could lead to beneficial results is provided in Sect. 11.4. Afterwards, Sect. 11.5 analyses the frictions and possible synergies today and in the future. Finally, concluding remarks and emerging trends for the immediate future are provided in Sect. 11.6.

11.2 Definitions and status quo of the IoT

This section defines the current state of the IoT with respect to the concept of the DE. The first subsection examines the different definitions of a 'Thing', adopted by standardization organizations, followed by our working definition for this chapter. The last subsection describes related works in which interaction between IoT and DE is the main goal.

11.2.1 One Concept, Many Definitions

The concept of a 'Thing' may seem generic. A 'Thing' can be characterized as a network object or entity that can connect to the Internet directly or through a network gateway. This exemplifies a network-centric perspective of the IoT in which a variety of interrelated 'Things' are able to communicate with each other to deliver new applications and services (Atzori et al. 2010). In contrast to the network-centric vision focusing on the communication technologies being used, the IoT can be seen from a purely Thing-centric perspective in which the services associated with Things are pivotal. These services are expected to manage large amounts of data captured by smart objects or 'Things' as a result of interacting with the environment.

Regardless of the vision, the definition of the term 'Thing' is extensive and includes a wide variety of physical elements. Examples of these elements include: (i) personal objects such as smartphones, smart watches or bands; (ii) ordinary objects and appliances in our daily lives such as refrigerators, lights, cars, and windows; (iii) other identifiable objects equipped with Radio-frequency identification (RFID) tags, Near-field communication (NFC), or Quick Response (QR) codes; and (iv) objects equipped with small microcontrollers.

Because of the heterogeneity of the technology and hardware, there is no single, unified definition of the term 'Thing'. Different international standardization bodies and organizations have suggested a definition, resulting in multiple interpretations of the concepts of Things and the IoT, which sometimes differ only slightly. Consequently, each stakeholder group may have a particular view of what the IoT and Things are, as demonstrated below by the definitions of some internationally renowned organizations.

The World Wide Web Consortium (W3C), an international organization whose aim is the collaborative development of Web standards, defines a 'Thing' as '*the abstraction of a physical or virtual entity that needs to be represented in IoT applications. This entity can be a device, a logical component of a device, a local hardware component, or even a logical entity such as a location* (*e.g., room or building*)' (Kajimoto et al. 2017).

The Institute of Electrical and Electronics Engineers (IEEE), a global professional engineering organization whose mission is to foster technological innovations and excellence for the benefit of humanity, defines a 'Thing' as a device with programmable capabilities. In contrast to the W3C's definition, the IEEE's definition takes a more practical engineering view of Things, driven by two defining features: (i) Things have the ability to communicate technologically, and (ii) Things have the ability to connect to or integrate in an already connected environment. This networking capability can be based on microcontrollers such as Arduino, Raspberry Pi, BeagleBone and PCDuino, among others.

The European Research Cluster on the Internet of Things (IERC) describes Things as '*physical and virtual things with identities, physical attributes, and virtual personalities and smart user interfaces, and are seamlessly integrated into the information network.*' (IERC 2014). Similarly, considering that Things belong to a network, the ITU introduces the term *infrastructure* and defines the IoT as "*a global infrastructure for the information society, enabling advanced services by interconnecting* (*physical and virtual*) *things based on existing and evolving interoperable information and communication technologies*" (ITU-T 2012). In addition, the ITU recognizes three interdependent dimensions that characterize Things (Fig. 11.1). This indicates the versatility of the IoT in application domains that differ in terms of the requirements and user needs.

The Internet Engineering Task Force (IETF), an open international community of network designers, researchers, and operators concerned with the evolution of the

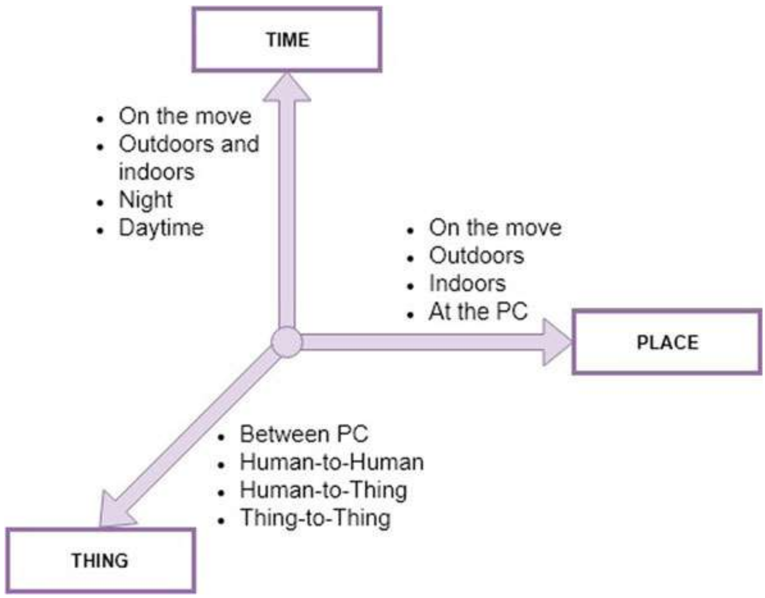


Fig. 11.1 Dimensions of the IoT (inspired in ITU-T 2012)

IoT, takes a broad perspective of Things in the context of the IoT, contemplating that "*'things' are very varied such as computers, sensors, people, actuators, refrigerators, TVs, vehicles, mobile phones, clothes, food, medicines, books,* etc. *These things are classified into three scopes: people, machines* (*for example, sensor, actuator,* etc.) *and information* (*for example, clothes, food, medicine, books,* etc.). *These 'things' should be identified at least by one unique way of identification for the capability of addressing and communicating with each other and verifying their identities. In here, if the 'thing' is identified, we call it the 'object'*" (Minerva et al. 2015).

Finally, the Organisation for the Advancement of Structured Information Standards (OASIS), a nonprofit consortium that drives the development, convergence and adoption of open standards for the global information society, describes the IoT as a '*system where the Internet is connected to the physical world* via *ubiquitous sensors*' (Cosgrove-Sacks 2014). OASIS focuses on the ubiquity of sensors, as they exist in '*every mobile, every auto, every door, every room, every part, on every parts list, every sensor in every device in every bed, chair or bracelet in every home, office, building or hospital room in every city and village on Earth*'.

In Fig. 11.2 we categorize the aforementioned IoT definitions based on physical, virtual and location considerations. The definitions reveal that these institutions and organizations consider the IoT from a physical point of view. In addition to the physical view, three organizations (ITU, IERC and W3C) add a virtual connotation to the definition of a 'Thing'. Only the W3C definition acknowledges explicitly location as a defining element of the IoT.

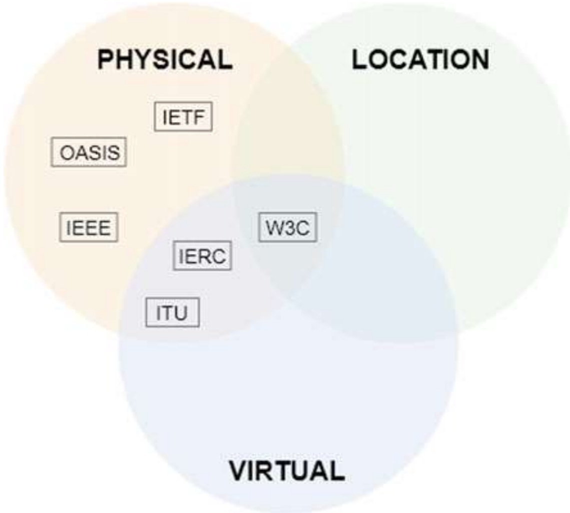


Fig. 11.2 Classification of IoT definitions

11.2.2 Our Definition

After analyzing the different definitions of internationally renowned institutions and standardization organizations, we propose our interpretation of the term ‘Thing’ that will be used throughout the rest of the chapter. This definition aims to (i) relate the IoT to DE, and (ii) be as broad as possible.

From our perspective, three main features characterize a ‘Thing’: (i) networked communication; (ii) programmability (data processing and storage); and (iii) sensing and/or actuating capabilities. From a DE perspective, the third feature plays a more prominent role. The sensing and/or actuating capabilities permit an IoT device or node to interact with its environment. This environment is closely related to the location feature, since all Things will intrinsically have this feature as a property, which increases in importance when the ‘Thing’ has a mobile component. Contrary to most of the definitions above, we consider a Thing’s location as a crucial characteristic because it impacts how a ‘Thing’ can communicate and how it can interact with its environment. However, we argue that the physical point of view can be understood to include location implicitly, as a physical sensor is located somewhere in the physical world.

11.2.3 Early Works on the Interplay Between DE and the IoT

As noted above, this chapter explores potential bridges between the IoT and DE for the development of applications and services that take advantage of the benefits of both infrastructures to effectively address complex research issues. In this context, we briefly summarize studies related to this objective.

In 1999, Gross predicted that electronic devices would populate the Earth and have the ability to capture different types of information, forming an 'electronic skin' (Gross 1999). These devices would be able to communicate through the Internet, and include meteorological or pollution sensors, cameras, blood pressure sensors or microphones, among others. The imagined 'electronic skin' could be in contact with what was happening in different scenarios and places on Earth, in the atmosphere, cities, houses, or even in ourselves.

Gross' vision is gradually becoming a reality. There is great variability in the form, size and purpose of sensors in wireless networks. Such Wireless Sensor Networks (WSN) enable distributed communication and data sharing between sensor network nodes. From this perspective, WSN form a subset of the IoT and, as such, the IoT can be seen as the logical next step of WSN in a progression that is still evolving in terms of the sophistication, variability in functionality, flexibility and integration with other infrastructures and network protocols (e.g., the Internet Protocol).

The IoT gained popularity between 2008 and 2013 (Fig. 11.3), and all organizations concerned with WSN began to focus on the IoT. The matured technology of WSN was applied to IoT developments, and DE organizations were not an exception. The field of sensors and sensor networks has been the object of study from multiple and varied angles, including the geospatial community, especially the Open Geospatial Consortium (OGC). The OGC started to transfer improvements made in the definition and application of standards and specifications in the field of WSN to the IoT.

The most significant OGC contribution concerning sensors and WSN has been the Sensor Web Enablement (SWE) standards suite (see Sect. 11.2.4 below). SWE enables the discovery and access of sensors and associated observational data through standard protocols and application programming interfaces (API) (Botts et al. 2008). The SWE standards have been applied directly to many application domains in DE.

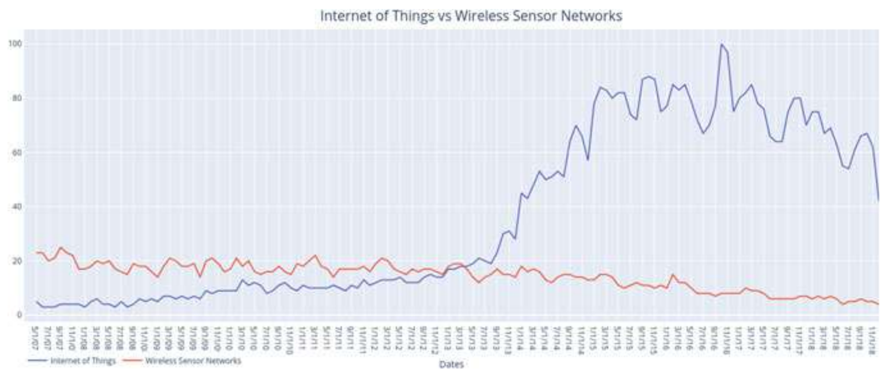


Fig. 11.3 Search volume on wireless sensor networks (red) and the Internet of Things (blue). *Source* Google Trends

The shared goal was to observe a particular phenomenon, for example, to predict emergency warnings or fire alarms or alerts when an event is triggered (Wang and Yuan 2010). For example, SWE has been widely applied to different Earth Observation (EO) application domains, with disaster management being one of the most important and well-developed. One of the early applications was the use of sensor web techniques to monitor natural and man-made hazards such as fires (Trilles et al. 2014; Jirka et al. 2009; Brakenridge et al. 2003), floods (Brakenridge et al. 2003), and volcanic eruptions (Song et al. 2008).

In parallel with the concept of WSN, Ashton (2009) noted that the term IoT was first used in his work entitled "I made at Procter & Gamble" in 1999. Back then, the IoT was associated with the use of RFID technology. However, the term WSN was not yet the focus of much interest, as shown in Fig. 11.3.

Some studies explored the connection between the IoT and DE concepts. Li and his colleagues studied the impact of the IoT on DE and analyzed the transition to Smart Earth (Li et al. 2014). The concept was introduced in 2009 during a panel discussion with the U.S. president and U.S. business leaders. In that panel, IBM's CEO Sam Palmisano requested that countries should invest in a new generation of smart infrastructure, with crucial use of sensors, suggesting the concept of 'Smart Earth' as a name. Subsequent governments showed interest in adopting this type of technology, and are making huge investments in researching and developing smart devices (e.g., the 'Array of Things' in Chicago, https://arrayofthings.github.io).

The primary objective of a 'Smart Earth' is to make full use of ICT and the IoT, and apply them in different fields (Bakker and Ritts 2018). In a 'Smart Earth', IoT devices are placed in all possible locations of our daily life, as long as our privacy can be respected. Through the combination of the IoT, DE, and cloud computing, globally deployed physical objects and sensors can be accessible online. The idea of a 'Smart Earth' is ambitious and includes remote sensing, GIS and network technology in combination with DE platforms (see Chap. 2 in this book featuring "Digital Earth Platforms"). The goal is to enable sustainable social development, which is a visionary step that is still utopian today, towards the establishment of a global information infrastructure to support UN Sustainable Development Goals (see Chap. 13 "Digital Earth for Sustainable Development Goals in this book,").

The work by Van der Zee and Scholten (2014) highlighted the importance of location in the concept of the IoT. The authors noted that space and time can play a role as 'glue', to enable an efficient connection between smart devices; therefore, geospatial sciences should have an active presence in the development of IoT architecture. In their study, Van der Zee and Scholten described a set of technologies related to the geospatial domain and big data analysis that could be combined with the IoT. The authors concluded that these technologies were already available for application in the field of the IoT and recommended their immediate use. However, the authors also identified the lack of IT professionals with knowledge in geospatial sciences as the main obstacle in massive uptake of the IoT for geo-related applications. They proposed to address this limitation through a gradual incorporation of core geospatial skills and competences into IT curricula.

Our aim in this chapter is to move beyond the initial steps and thoughts presented in Van der Zee and Scholten (2014), where the status quo of the IoT and DE was described five years ago. We focus on the 'current status quo' by outlining emerging technology trends that can be crucial for establishing real connections between DE and the IoT, and investigate developments during the last five years in particular. Even though development has been gradual and incremental, and not rapid and revolutionary (i.e. from a GIScience perspective), new requirements and technology trends have appeared and the IoT has become a topic that is undoubtedly gaining increasing traction.

11.2.4 IoT Standards Initiatives from DE

As noted above, the IoT ecosystem has been very diverse for several years (Atzori et al. 2010), and its diversity has been increasing. It is comprised of heterogeneous devices, protocols and architectural approaches. A plethora of international initiatives are put in place to unify and streamline aspects associated with the design and implementation of IoT infrastructures. The current standardization initiatives address aspects related to discoverability, data transmission, device processing and tasking.

The growing number of interconnected devices, combined with the increasing importance of the use of the IoT in almost any aspect of human life, tend to increase the need and importance of mature, well-established and -implemented standards. The diversity of different standardization initiatives provides designers and developers with a broad range of opportunities that do not necessarily complement each other. There are multiple ways of reaching the same destination, i.e., there is no single solution to be adopted. Here, we provide a short overview of selected IoT standards that play an important role within the context of DE. The SWE suite of standards is described in more detail in Chap. 8 of this book.

From the geospatial perspective, the OGC coordinates different standardization initiatives. This consortium is comprised of more than 525-member organizations from governmental, commercial, non-governmental, academic and research institutions. The primary objective of the OGC is to develop open standards that include a geospatial component. These standards are developed through a consensus-based process and are openly available to streamline the exchange of geospatial data. OGC standards are used in a wide variety of domains, including geosciences and the environment, defense and intelligence, emergency and disaster management, and public services, among others.

Over a decade ago, well before the IoT became mainstream, the OGC developed the SWE suite of standards for spatio-temporal observation data (Botts et al. 2008). SWE outlines a set of specifications related to sensors and proposes data models and Web service interfaces that can act as a bridge between sensors and users, allowing the sensors and their measurements to be accessible and controllable through the Web (Sheth 2018). The SWE suite, although initially designed for sensors, can easily be applied to any type of spatio-temporal data flow (including heterogeneous types of smart devices with an observation capability). It offers a set of specifications in an open standard schema using extensible markup language (XML) and web services. It enables (i) finding sensors and sensor data; (ii) describing sensor systems and data; (iii) recovering real-time and historical sensor observations; (iv) adding simulations and recovering simulation results; (v) reporting results and alerts; and (vi) full web control.

SWE (depicted in Fig. 11.4) is organized through several interdependent standards that include the Sensor Model Language (SensorML) (Botts and Robin 2007), Observations and Measurements (O&M) (Cox 2003), Sensor Observation Service (SOS), Transducer Markup Language (TransducerML, deprecated) (Havens 2007), Sensor Planning Service (SPS) (Simonis 2007), Sensor Alert Service (SAS) (Simonis 2006) and Sensor Event Service (SES) (Echterhoff and Everding 2008). In this work, only the first three specifications are shown in detail (i.e. SensorML, O&M, SOS), as they are the most widely used in the IoT context today.

SensorML provides the ability to define a sensor in a structured manner. The standard specifies how to find, process and record sensor observations so that a data model and XML schema can be established to control sensors through the Web. SensorML defines a standard schema describing any type of sensor, stationary or dynamic, in situ or remote, active or passive. The PUCK protocol (O'Reilly 2010) is an addition to the SensorML standard that provides a low-level protocol to retrieve sensor drivers, and metadata documents, encoded according to SensorML.

The O&M standard, initially developed by the OGC, is also adopted as an International Organization for Standardization (ISO) standard (ISO 2011). It provides a model for representing and exchanging sensor observations. The standard is encoded

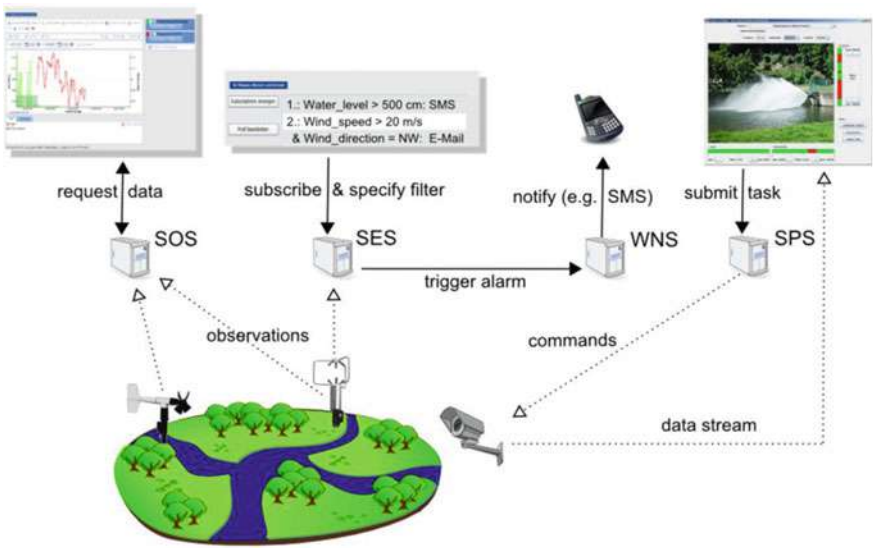


Fig. 11.4 The sensor web enablement suite of standards. *Source* Bröring et al. (2011)

using an XML/JSON data model, which describes the relationship between different aspects of the data capture process. The O&M schema defines both observations and phenomena. In addition, it can be extended to better support metadata.

Finally, the SOS provides an interoperable means for serving observations via a Web interface and is the primary service model of the SWE suite. The current version of the standard introduces a modular structure. The base module provides three mandatory operations. The first, "GetCapabilities", offers a spatial and temporal description of the observations that have been stored, as well as a list of the sensors and their available features. The "DescribeSensor" operation is used to return a sensor description using SensorML. The "GetObservation" operation provides access to the actual spatio-temporal data encoded in accordance with the O&M standard.

All the standards described above were conceptualized and adopted several years ago within a completely different technological landscape. The rapid growth of the IoT and the emergence of new technologies (e.g. remote sensing, 4G/5G communication, machine-to-machine and machine-to-human interactions) brought new challenges such as (i) the need for lightweight data encoding, (ii) the need for higher bandwidth for data exchange, and (iii) the issue of constrained devices with little or no computational capabilities, such as RFID tags and QR codes (Kotsev et al. 2018). These challenges acted as a driver for the OGC and led to adoption of new standards that better fit the IoT.

The SensorThings API (Liang et al. 2016), designed to follow the paradigm of the Web of Things (WoT) (Guinard et al. 2010), offers access to data through standard web protocols and is based on the O&M conceptual data model. The main features of the standard are (i) a RESTful interface, (ii) the use of lightweight and efficient JSON encoding, (iii) adoption of the OASIS OData URL pattern (OData) and query options, and (iv) support for the ISO message queuing telemetry transport (MQTT) messaging protocol to offer real-time connections.

The SensorThings API data model (shown in Fig. 11.5) is divided into two parts (profiles), namely, the 'Sensing' profile and the 'Tasking' profile. The former enables IoT devices and applications to CREATE, READ, UPDATE, and DELETE (through the standard web operations HTTP POST, GET, PATCH, and DELETE) IoT data and metadata by invoking a SensorThings API service. In addition, the tasking profile provides a standardized approach for controlling IoT devices through the "ACT" capability, which is revisited in the next section. Each 'Thing' has a Location (or some Historical Locations) in space and time. A collection of Observations grouped by the same Observed Property and Sensor is called a Datastream. An Observation is an event performed by a Sensor that produces a value of an Observed Property of the Feature of Interest.

From a spatial analysis perspective (De Smith et al. 2018), many raster- and vector-based operators and techniques have been developed over the last decades and have been shown to be successful in many varied applications. Substantial progress has been made to bring geospatial workflows—i.e., a combination of the above spatial operations to accomplish a sophisticated analytical process—to the cloud and distributed computing environments (e.g., Granell et al. 2010; Granell 2014; Yue et al. 2016), expanding the field of the Geoprocessing Web (Zhao et al. 2012) to the

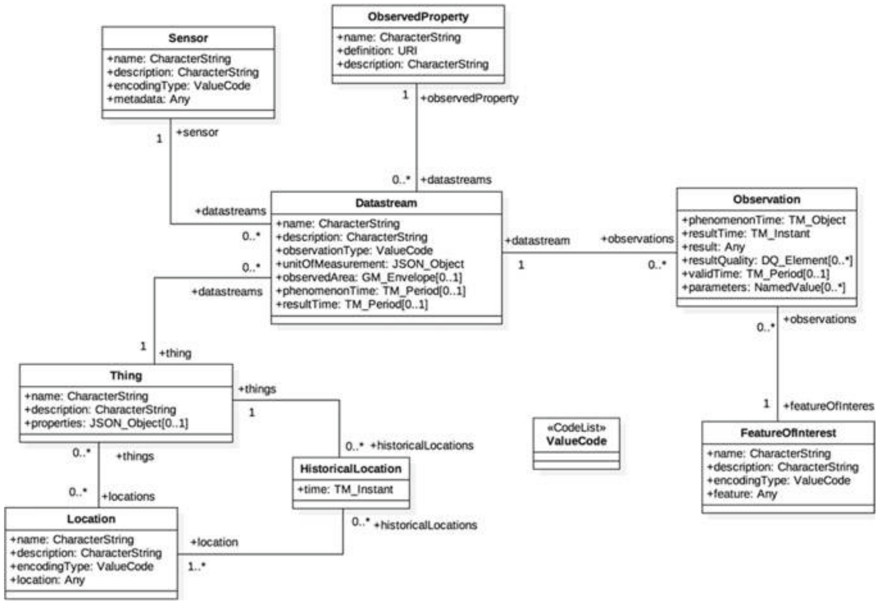


Fig. 11.5 The SensorThings API data model. Each thing has a location (or some historical locations) in space and time. A collection of observations grouped by the same observed property and sensor is called a datastream. An observation is an event performed by a sensor that produces a value of an observed property of the feature of interest. *Source* OGC SensorThings API (http://docs.opengeospatial.org/is/15-078r6/15-078r6.html)

Digital Earth (Hofer et al. 2018). The OGC Web Processing Service (WPS) (OGC 2005), a service interface for exposing and executing processes of any granularity on the Web, enables sharing and integration of spatial data processing capabilities on the Web, including polygon area calculation, routing services, or entire environmental models (e.g., Díaz et al. 2008; Granell et al. 2010). The geoprocessing capabilities in DE are extensively covered in other chapters, e.g., Chap. 5, and our interest lies solely in the relationship between the WPS and the IoT (see Sect. 11.3.2).

11.3 Interplay Between the IoT and DE

One of the aims of this chapter is the identification of potential bridges between the IoT and DE. This overview is partly speculative since we tried to identify potential paths for collaboration between both infrastructures, which may or may not lead to successful linkages in the future. To support our claims in Sect. 11.4, we identify the current situation, i.e., the state of the art of the IoT's and DE's technological substrate. In this section, we highlight new technological developments and emerging trends that are or may become crucial in the coming years that were not present or not sufficiently developed at the time of Van der Zee and Scholten (2014).

Along the lines of the topics described in Sect. 11.2.3, the traditional focus of DE embraces the following high-level functions (Lü et al. 2019): (i) discovery and acquisition of spatial information, (ii) understanding of spatial objects and their relationships (e.g., GIS analysis, spatial statistics), and (iii) determination of the spatio-temporal behavior and simulation rules (e.g., simulations, predictions). These functions help categorize and restrict the discussion in terms of the current technological substrate. However, we should interpret and contextualize these high-level functions of DE from the viewpoint of the IoT.

First, the acquisition of spatial information is a crucial function in the IoT because Things and smart devices observe and sense their environment to collect observational measurements. Through the lens of the IoT, the discoverability of Things and the communication of gathered spatial data become extremely relevant for data acquisition. Of the two main capabilities of Things (see Sect. 11.2.2), the ability to **observe and sense**, is a fundamental mechanism to provide input observational data for DE.

Second, spatial statistics and spatial analysis are well-established geospatial methods for exploring spatial patterns, relationships and distributions (De Smith et al. 2018; Worboys and Duckham 2004). Analytical methods are fundamental building blocks in DE, although recent trends in real-time analysis and edge computing promise to move much of the analytical power to devices (i.e., edge and fog computing) so that gathered data can be immediately processed directly on the smart devices. This trend suggests that analytical improvements in the IoT will also play an important role in DE.

Third, predictive modeling and simulations are required to explore both physical and social dynamic geographic phenomena to better understand the evolution, changes and dynamics of the phenomena from a spatio-temporal perspective, to gain new insights and scientific knowledge to support informed decision-making processes. Understanding spatiotemporal behaviors makes sense from the DE point of view, to aid in the assembly of a detailed yet broad perspective of the complex, multidimensional relationships that occur in the real world. We recognize that prediction and simulation activities are typically associated with DE and that advances in the IoT might contribute to this area, but we see this hypothetical scenario occurring in the mid- to long-term, well beyond the time frame of the speculative exercise in Sect. 11.4. Since research on the IoT and DE with respect to predictive modeling and simulations is still in its infancy, we do not cover it in this chapter.

As a result of the previous functions, new scientific knowledge is generated that is necessary for taking informed and insightful actions, often ‘acting’ over the environment. In terms of **acting**, the second main capability of Things, new knowledge can trigger actions at least at two different levels in the context of the IoT: first, self-calibration of a sensor and/or Thing, similar to adjusting the lens in a human eye to sharpen the image, e.g., changing the sampling frequency; and second, providing a reflex similar to a reaction to pain without thinking, e.g., by opening a valve or level in the case of imminent flooding. However, this view would mean a priori that

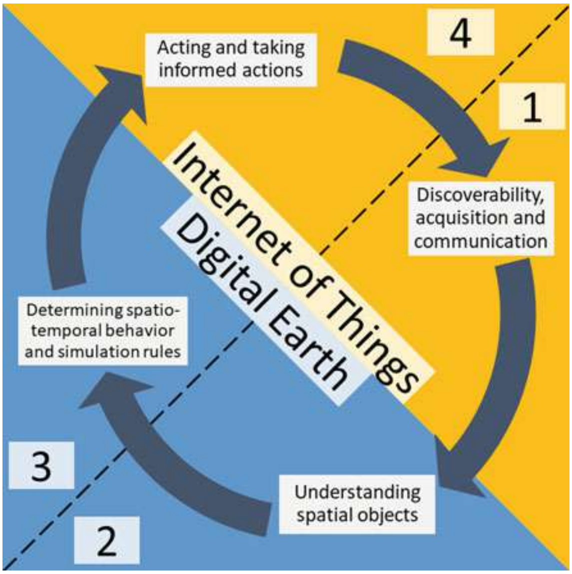


Fig. 11.6 IoT and DE workflow according to the higher cognitive functions in DE

the acting in IoT and Things do not contribute sufficiently to the higher (cognitive) functions of DE such as spatial analysis, predictive modeling and simulation, but the results of higher cognitive functions in DE may impact the acting behavior of Things and the IoT. In addition, we add a fourth function related to the ability of Things to act and take informed actions, depending on the insights and knowledge produced in the analysis, simulations, and predictions in DE.

Figure 11.6 reflects the existing and potential roles of each infrastructure in relation to the four functions: (i) discoverability, acquisition, and communication of spatial information, (ii) understanding of spatial objects and their relationships, (iii) determining spatio-temporal behavior and simulation rules, and (iv) acting and taking informed actions. We argue that the IoT infrastructure is important in (i) and (iv) whereas DE is more relevant in (ii) and (iii). For (i), the IoT can enhance DE by acquiring data streams from new sources, at a fine scale and high frequency. For (ii), it is plausible that both infrastructures progressively collaborate in a symbiotic manner per use case. From a broader perspective, it can reasonably be argued that DE includes IoT and encompasses the IoT life cycle in a broader ecosystem. Although GIS methods and analysis have traditionally taken a predominant role in DE, the role of the IoT will most likely increase in the future given the close relation between the IoT and the nascent edge-fog-cloud computational paradigms that enable IoT-based analytical processes to be conducted at different scales. This is a partial view, as we focus on the relationship between DE and the IoT. For example, remote-sensing satellite imagery, LIDAR and UAV were intentionally omitted even though they are key spatial data sources (i.e., the first function) for DE. We acknowledge the fuzziness of the boundary between both infrastructures and pay special attention to the interplay between DE and the IoT in Fig. 11.6, demonstrating how collaboration and integration is starting to happen while frictions and barriers are becoming more visible.

In the following sections, we identify for all but the third function the current technological substrate.

11.3.1 Discoverability, Acquisition and Communication of Spatial Information

Discoverability of Things. An important objective in IoT research is the discovery of devices and their services and/or the data they produce. The absence of standardized discovery methods for the WoT (Zhou et al. 2016) led to the development of online global sensor directories and collections such as Xively (https://xively.com), SenseWeb (Grosky et al. 2007), SemSOS (Pschorr et al. 2010) and the SWE discovery framework (Jirka et al. 2009). A key feature of these online directories/registries is that they provide open Web APIs supporting the development of third-party applications. The main drawback is that they are centralized, with a single point of failure. Decentralized approaches have also been proposed, such as IrisNet (Gibbons et al. 2003), which uses a hierarchical architecture for a worldwide sensor Web. G-Sense (Perez et al. 2010) is a peer-to-peer (P2P) system for global sensing and monitoring. These approaches, although more robust and scalable, do not effectively solve the problem of sensor discovery as they still require sensor registration to dedicated gateways and servers, which need to maintain a hierarchical or P2P structure among them.

Approaches towards real-time discovery of physical entities include Snoogle (Wang et al. 2008) and Dyser (Elahi et al. 2009). Snoogle is an information retrieval system for WSNs, but it cannot scale for the World Wide Web. Dyser requires an additional Internet infrastructure such as sensor gateways to work. Moreover, utilization of the domain name system (DNS) as a scalable, pervasive, global metadata repository for embedded devices and its extension for supporting location-based discovery of Web-enabled physical entities were proposed (Kamilaris et al. 2014; Kamilaris and Pitsillides 2012). However, this technique requires changes in the existing Internet infrastructure. It is possible to exploit web crawling for discovery of linked data endpoints, and through them the discovery of WoT devices and services was examined in WOTS2E (Kamilaris et al. 2016) as well as in SPITFIRE (Pfisterer et al. 2011).

While the approaches described above are mainly targeted at 'professional' users, there is demand for a simple and easy means for the general public to access IoT data. Experts can use a plethora of different service interfaces and tools to discover and utilize data from IoT devices, as implemented by the SmartEmissions platform (Grothe et al. 2016). Nonexpert users typically only search for IoT devices and their data through mainstream search engines such as Google and Bing. Ensuring the discoverability of devices and the data they produce is being investigated for geospatial data in general (see Portele et al. 2016 for further details). A similar approach might be adopted for the IoT, considering its higher complexity due to the high temporal (and spatial) resolution of the data produced by Things.

Spatial acquisition with Things. Some examples of geospatial standards to encode sensor metadata and observations were introduced in Sect. 11.2.4, and the SensorML standard is one of the most important. SensorML describes sensor metadata in a comprehensive way, providing a useful mechanism to discover sensors and associated observations. This standard specifies information about a sensor such as its sensor operator, tasking services, location, phenomena, and history of the sensor. Thus, it can be used by discovery services to fill their search indexes.

Following the SWE framework, there are two different search types (Jirka et al. 2009): *sensor instance discovery* and *sensor service discovery*. The first type finds individual sensors (devices) or sensor networks, and the second type refers to services that interact with the sensor (through sensing or tasking). Jirka et al. (2009) define three different criteria to identify both annotated search types:

- The Thematic criterion covers the kind of phenomena that a sensor observes, such as temperature, humidity, or rainfall.
- The Spatial criterion refers to the location where the sensor is deployed.
- The Temporal criterion is the time period during which the observations are generated.

This classification was defined from a conventional sensor point of view. The inclusion of current IoT devices with the ability to act leaves the previous criteria incomplete, as some IoT devices act as well as observe. Therefore, the definition of the thematic criterion requires extension to include an IoT device's capability to act, for example, to turn on/off a light or activate/deactivate an air conditioner.

In addition to the three shared criteria, Jirka et al. (2009) defined two criteria that focused exclusively on the *sensor instance discovery* type of search: sensor properties and sensor identification. The sensor properties are based on a specific state of the sensor, for example to find all *online* sensors. The sensor identification refers to the unique id used to identify unambiguously a sensor. Regarding the *sensor service discovery type of search*, two additional criteria were defined: functionality and usage restrictions. The first refers to the functionalities of the associate service such as available operations for data access, alerting or tasking, among others. The second criterion on usage restrictions is related to the permissions and restrictions to access the service functionalities.

Two different aspects are vital for the successful discovery of a sensor: metadata and semantics. As for all spatial data, metadata is essential to describe and discover a sensor or a network of sensors. SensorML was created for this purpose and can define a sensor in a well-known manner to add flexibility and allow for the use of any type of sensor. The Sensor Instance Registry (SIR) defines operations for handling sensor metadata and allows for sensor discovery. The above criteria, both common and specific for each type of search, are closely related to the metadata aspect for the discovery of sensor instances and services.

Semantics is the other pillar in a powerful and effective discovery service. Semantic rules can aid in locating sensors related to the same phenomena or discovery of all sensors that are related to the same thematic aspect. This semantic view can be extrapolated to link sensors with places to retrieve sensors or observations associated with place names. The Sensor Observable Registry (SOR) offers a primary interface to explore this kind of relationship between phenomena and sensors.

Unfortunately, the support of semantics is a weakness in the SWE standards. To solve this issue, an initiative from World Wide Web Consortium (W3C) was created to integrate and align sensors with semantic web technologies and Linked Data. This contribution was led by the W3C Semantic Sensor Network Incubator Group (SSN-XG) that proposed an ontology called Semantic Sensor Network (SSN) to address the semantic gap in sensor-related OGC standards (Compton et al. 2012). The main fields of this ontology are sensors (e.g., location, type), properties (e.g., precision, resolution, and unit), and measurements (values).

Despite the great advances that SSN brought, it does not currently support all the possibilities that the IoT offers since SSN was designed before the mainstream adoption of the IoT. New ontologies have been launched to cover this gap. One example is how the Internet of Things Ontology (IoT-O). IoT-O adds some missing concepts relevant to the IoT such as Thing, Actuator, and Actuation (Seydoux et al. 2016). Similarity, the Sensor, Observation, Sample, and Actuator (SOSA) ontology is a follow-up to SSN. It is the result of a joint effort of the W3C and OGC that builds on the lessons learned from SSN to provide a better representation of the IoT and alignment with OGC-related specifications (Janowicz et al. 2018).

Communication with Things. The advances in IoT connectivity solutions such as Bluetooth, ZigBee, Wi-Fi and 3-5G (Palatella et al. 2016) combined with decreases in the price and energy consumption of IoT components have led to a huge deployment of smart devices using IP-connectivity worldwide, increasing the frequency of communication to the point that they are perceived as always connected. As outlined above, these devices can offer two different capabilities, observing (sensing) and acting. A decade ago, sensor networks were only able to capture and send data, similar to a simple data logger. In recent years, the ability to establish two-way communication between Things and the cloud has added the feature that Things can (re)act. Consequently, new protocols that enable machine-to-machine (M2 M) communication have been developed, with the goal of providing efficient and transparent two-way communication channels between smart devices. Examples of such TCP/IP-based protocols are the advanced message queuing protocol (AMQP), MQTT, and the simple/streaming text oriented messaging protocol (STOMP). These communication protocols are adapted to the requirements of IoT devices that are constrained concerning their performance and energy efficiency.

11.3.2 Spatial Understanding of Objects and Their Relationships

Spatial analysis of Things. There are many more smart devices (Things) around today than five years ago. Smart devices now produce massive volumes of data, i.e., flows of data with strong temporal and spatial features. Therefore, spatial analytical methods such as proximity, area, volume, and trajectory are of vital importance in analyzing processes of Things. However, the variety of data sources related to the IoT has posed new analytical challenges, especially in the design and provision of a new class of analytical tools capable of handling real-time temporally and spatially referenced data from a plethora of heterogeneous smart devices (Trilles et al. 2017). Despite the existence of tools capable of analyzing temporal data in real time, the same does not appear to be true for the spatial component. Space (location and orientation for all Things, size and shape for larger Things such as cars) plays an indispensable role in the IoT, as Things-generated data have spatial properties and are spatially related to each other. Promising initiatives and platforms have recently emerged with the aim of performing spatio-temporal analysis in real-time, such as Microsoft Streaminsight, the Oracle Spatial Database with the Oracle Complex Event Processing engine, and the GeoEvent processor module as an extension of the ArcGIS Server environment (ArcGIS Server, n.d.).

Despite these notable efforts, spatial support for the real-time analysis of IoT data is still in its infancy. As Van der Zee and Scholten (2014) noted, any IoT architecture should consider the geospatial component. Location provides a kind of 'glue' that efficiently connects smart devices. The authors proposed storing the location of each 'Thing' and other geographic-related features such as orientation, size, and shape. However, the ability to handle and analyze the location of Things in near real time is still limited with existing analytical platforms, despite its opportunities (McCullough et al. 2011; Rodríguez-Pupo et al. 2017).

Furthermore, spatio-temporally located Things have the potential to significantly improve advanced geospatial analysis, as Kamilaris and Ostermann (2018) describe in their review on the potential role of geospatial analysis in the IoT field. In short, Kamilaris and Ostermann suggest network analysis and monitoring, surface interpolation, and data mining and clustering as spatial analysis techniques and methods that would especially benefit from an increasing number of mobile or stationary sensor Things. However, as the authors noted, these advanced analytical applications have been scarcely exploited to date.

Geospatial standards for Things. Despite some remarkable exceptions such as prototype systems to analyze data from air quality sensor networks (Trilles et al. 2015b), real-time, geospatial analysis approaches and tools have not been sufficiently developed to offer standardized procedures through uniform interfaces that can be widely consumed and integrated in DE applications. DE has traditionally considered sensors as a fundamental pillar to collect information to support and realize strategies or policies at a higher level. As described in Sect. 11.2, the SWE suite was the initial step in offering a standardized specification that would fulfil the requirements demanded by the IoT from the DE perspective. For example, the SOS specification requires handling large XML documents, which is problematic in a typical scenario in the IoT where memory capacity and connectivity are limiting factors.

Although the core of the SWE suite has served to cover the required functionality of the IoT, the complexity of the data models in some of the specifications (Tamayo et al. 2011; Trilles et al. 2014) and the appearance of new requirements such as the ability to work in real time and to act have reduced the applicability and integration of the SWE suite in the scope of the IoT. In an effort to bridge the gaps between SOS and the IoT, new extensions or approaches attempt to make the SOS interfaces more suitable for IoT devices. These approaches include SOSLite (Pradilla et al. 2015), TinySOS (Jazayeri et al. 2012) and SOS over CoAP (Pradilla et al. 2016).

Another crucial feature for the analysis functionality of the IoT and Things is the ability to specify and perform real-time and asynchronous notifications and communications. In this regard, the GeoMQTT protocol based on the MQTT protocol allows for adding spatial notification and data streaming between publish/subscribe instances (Herle and Blankenbach 2018). Following the original approach of the MQTT channels, the authors proposed the concept of GeoPipes to distribute instances and enable the sharing of geospatial data streams in a standardized manner.

Laska et al. (2018) proposed a real-time stream processing pipeline that allows for spatiotemporal data stream integration from IoT devices. A data integration layer allows for geospatial subscriptions using the GeoMQTT. Tools such as Apache Kafka and Storm are used to transfer and apply map matching algorithms to IoT data with spatiotemporal components. For example, these algorithms were used to analyze traffic congestion for a recent route optimization using IoT Things with Global Navigation Satellite System (GNSS) receivers in buses.

Another study (Rieke et al. 2018) took an additional step to bridge the DE and IoT realms by arguing for the need to establish event-driven architectures as a natural evolution of the predominantly static Spatial Data Infrastructures (SDI). The authors identify a series of interdependent issues that need to be addressed in the coming years to take full advantage of the uptake of eventing in GIScience (and DE). The issues relate to the (i) inconsistencies between classic data access methods that are based on a request-response pattern, and event-driven approaches where a publish-subscribe pattern prevails, (ii) heterogeneous approaches for defining event patterns, (iii) multiple standards and limited support in software tools, (iv) the integration of devices in an SDI and the data they produce, and (v) the lack of semantic interoperability of geospatial events.

11.3.3 *Taking Informed Actions and Acting Over the Environment (ACT)*

As shown in the defined IoT lifecycle (Fig. 11.6), to act means to take or perform actions (over the environment) depending on the results obtained in previous functions. Bélissent (2010) noted that this feature can make the management of public services in a city, education, health, safety, mobility or disaster management more aware, interactive and efficient.

IoT devices have been traditionally suitable for use as input sources for Decision Support Systems (DSSs) in a multitude of application domains and use case scenarios such as disaster management, cities, mobility, and safety. In this chapter, we focus on Spatial Decision Support Systems (SDSSs), which are defined as interactive systems designed to support decision making related with spatial planning problems. SDSSs have evolved to more complex architectures and communication models, from systems deployed on the cloud operating with data from the WSN (or IoT data sources) to a shift in the computing paradigm in which the actual computation is implemented at three different levels: edge, fog, and cloud (Fig. 11.7). In this new setting, both the computation and decisions are made closer to the producers of the data (*Things*).

The 'Edge' is the layer that covers the smart devices and their users, providing local computing capacity within Things. The 'Fog' layer is hierarchical, aggregating a variable number of edge layers. In addition to computing, the fog layer has other functionalities such as networking, storage, control, and data processing, possibly using data produced by the edge layer and data from other sources. As a result, data contextualization is more important in the fog layer to make sense of different data sources than the typical single data stream in an edge layer. The 'Cloud' layer on top performs the final analysis to extract information and create knowledge to be

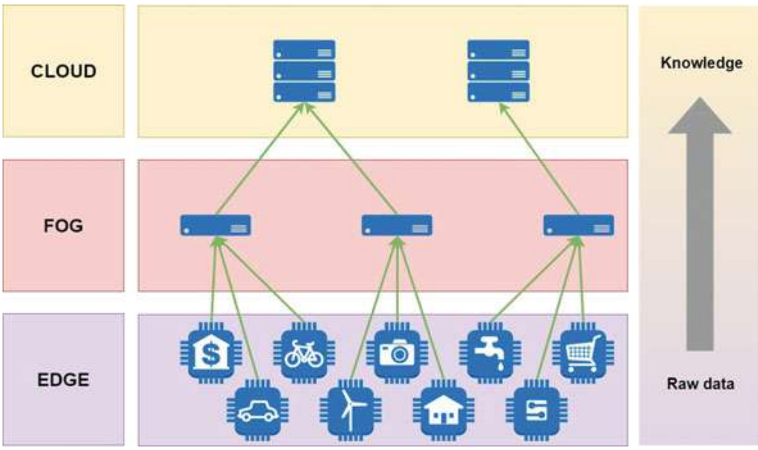


Fig. 11.7 Three-layer IoT architecture

transferred for decision support actions. This implies an increased level of contextualization and complexity in the analysis process than in the previous (lower) layers, at the cost of losing capacity for real-time analysis.

Given the edge-fog-cloud layered architecture, the introduction of geospatial concepts and spatial analysis in the fog layer could allow for decision-making processes without a human in the loop based entirely on the semantics of the spatial-temporal dimensions in the incoming data. In recent years, many efforts have been made to move the analysis from the cloud to the fog layer, with the aim of reducing latency in the analysis once the data are received in the fog layer (Barik et al. 2016).

Although data usually flow from the edge to the cloud layer (sensing capability), devices with the ability to act (tasking) also require information to perform their operations. The tasking capability allows for other devices or users to actuate devices via the Internet so that these 'controlling' devices or users can easily control them to execute tasks remotely. Autonomous Things would be previously programmed to act without establishing a connection. While the sensing capability allows for users to continuously monitor the status of devices and the environmental properties they capture, the tasking capability can help users make adjustments accordingly by controlling devices remotely.

In general, combining the sensing and tasking capabilities of IoT devices enables users to create various automatic and efficient tasks and applications. These kinds of applications are called "physical mashup" applications (Guinard et al. 2010). A simple, domestic example is the activation of an air conditioning system depending on the position and behavior of the user, through an application that uses a GNSS sensor. In this example, the air conditioning device provides an interface to turn on/off (tasking) the system to establish a comfortable temperature. To facilitate this kind of mashup of sensing and tasking capabilities, a uniform (interoperable) interface for users or applications to enable access and communication is a critical requirement.

The tasking feature was initially conceived in the SPS specification of the SWE suite. SPS offers a standardized interface for tasking sensors and sensor systems and defines interfaces to expose sensor observations and metadata. For example, a sensor network can be set up to measure air pollution in 5-min intervals or a satellite can be tasked to remotely sense a specific region on the surface of the globe (De Longueville et al. 2010). This standard offers operations such as GetFeasibility, which can be used in advance to verify whether the execution of a task is feasible for a certain sensor, and the DescribeResultAccess operation to determine the access points to collected data. The SPS interface also offers functionality for managing submitted tasks, including convenient operations for retrieving the status of a task, updating tasks or cancelling them.

A next step is the tasking profile of the SensorThing API, which is a follow-up, improved profile of the SPS (Simonis 2007). The SensorThing API (see Sect. 11.2) defines two different profiles, Sensing and Tasking. The Tasking profile is based on the SPS standard and enables interoperable submission of tasks to control sensors and actuators. The main difference between SPS and the SensorThings API is that the former offers task operations over sensors and the latter also includes tasks on actuators. Although the first version of the SensorThing API did not include the

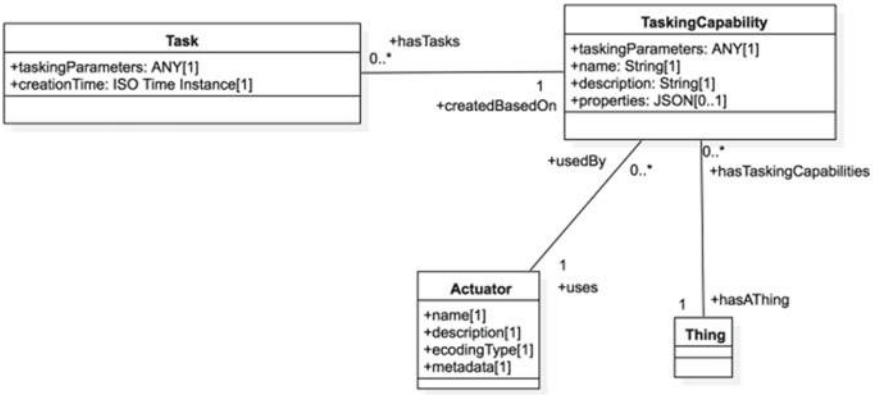


Fig. 11.8 The SensorThings API tasking entities. *Source* OGC SensorThings API (http://docs.opengeospatial.org/is/15-078r6/15-078r6.html)

Tasking profile, a new candidate standard illustrates the potential of the SPS standard, duly adopted and aligned with the requirements of the SensorThings specification (Liang and Khalafbeigi 2018). This new specification called Tasking Core defines three new entities, *TaskCapability*, *Task,* and *Actuator* (Fig. 11.8).

The *TaskingCapability* entity describes all supported tasks for each Thing and how they can be used. This entity is defined by four properties: name, description, taskingParameters, and properties. The second entity, *Task*, is a list of performed tasks that are defined by a set of tasking parameters (commands executed) and creation time. The last entity is the *Actuator* and defines a type of transducer that converts a signal to a real-world action or phenomenon. This entity is comprises a name, description, encoding type of metadata and metadata.

11.4 Case Studies on Smart Scenarios

In this section, we show how the IoT and DE work hand-in-hand in real-world scenarios based on the latest technology initiatives to relate the IoT and DE described in the previous section. Kamilaris and Ostermann (2018) provide an extensive overview of work at the nexus of geospatial analysis and the Internet of Things; here, we provide a selection of case studies in various domain applications, with a special focus on the relationship between DE and the IoT.

In the context of applications for environmental monitoring and resource management in cities, recent examples of IoT applications include an Arduino-based sensor platform in Seoul to measure variations in the physical-chemical parameters in water streams (Jo and Baloch 2017). The sensor platform is powered by solar energy and transmitted sensor readings every second via Bluetooth for three years. Although the case study in Jo and Baloch (2017) relies on a single sensor station and the clustering analysis of the raw data focuses uniquely on the temporal dimension, the paper shows the potential of Arduino-based sensing modules for environmental sensing applications in smart city applications. To improve solid waste management, Tao and Xiang (2010) developed an information platform to support recycling. The main technologies were RFID and GPS to track and check waste flows between collection, transport, and processing facilities. Lee et al. (2015) examined the role of the IoT in an industrial service provision scenario (fleet management) and Fazio and Puliafito (2015) use the example of road conditions to showcase a cloud-based architecture for sensor and data discovery. They distinguish two scenarios of data- or device-driven search, and develop the system architecture based on the OGC SWE suite and the extensible messaging and presence protocol (XMPP).

Reducing the required energy consumption remains an important objective for IoT devices. Ayele et al. (2018) proposed a dual radio approach for wildlife monitoring systems. They combine Bluetooth low energy for intraherd monitoring with LoRa for low-power wide-area networks to communicate between herd clusters and a monitoring server. The proposed architecture promises significant advantages in reducing power consumption while maintaining low latency.

Improving traffic management is another promising IoT application area. In 2006, Lee et al. proposed the use of cars as a mobile vehicular sensor network and for data exchange in "smart mobs". More recently envisioned solutions include parking management and smart traffic lights as part of a cognitive road management system that handles different types of traffic efficiently (Miz and Hahanov 2014). Jing et al. (2018) examined the combination of GNSS localization and RFID tagging for infrastructure asset management with promising results. Additionally, the city of Aarhus in Denmark deployed traffic sensors across major roads in the city, and the information was used by the CityPulse project to provide context-aware recommendations to users for route planning (Puiu et al. 2016).

Noise pollution is a frequent problem in dense urban areas, and because urban morphology makes noise distribution modeling difficult, it has attracted participatory sensing approaches. Wireless acoustic sensor networks are another option. Segura Garcia et al. (2016) presented a case study in the small city of Algemesi (Spain), where a network of 78 inexpensive sensor nodes based on Raspberry PIs collected sufficient data for a subsequent highly accurate spatial interpolation.

Okasenen et al. (2015) harnessed movement data from mobile sports tracking applications in urban areas to produce heat maps of cyclists commuting through the city of Helsinki. Mobile phones could be considered IoT sensor devices in participatory sensing-based models for mining spatial information of urban emergency events, as demonstrated by Xu et al. (2016). In addition, van Setten et al. (2004) supported the COMPASS tourist mobile application with context-aware recommendations and route planning. Mobile phones were also used for crowdsourcing-based disaster relief during the Haitian earthquake (Zook et al. 2010), where people used the camera and GPS of their phones to send information from the field to the authorities to map the landscape of the disaster and assess the overall damage.

University campuses present an interesting environment for smart city approaches because the visitors are usually more tech-savvy than the average population, the network coverage is good, and the geographic boundaries allow for a comparatively crisp delineation of the study area. Cecchinel et al. (2014) presented a system architecture for a smart campus case where the four requirements of sensor heterogeneity, reconfiguration capability, scalability, and data as a service were handled via a middleware in the Amazon Web Services (AWS) cloud, with Arduino Uno and Raspberry Pi sensors for bridging. Another case study at a university campus examined the impact of nearby weather and pollution sensors on the everyday decision-making of the students (Kamilaris and Pitsillides 2014). Trilles et al. (2015a) presented a sensorized platform proposal that adheres to the principles of the IoT and the WoT. They use the SensorThings API to avoid interoperability issues. An environmental WSN in a Smart Campus scenario was developed as a proof of concept.

However, smart approaches with IoT technology are not limited to smart city applications. Sawant et al. (2014) presented a low-cost automated weather station system for agriculture that uses Raspberry Pi systems at its core and SWE to transmit data. The sensor readings were also broadcast on a dedicated Twitter account. The system has been extended with additional components such as a web-based client (Sawant et al. 2017). The environmental impact of agriculture was studied by Kamilaris et al. (2018) in the region of Catalonia, Spain. In their study, sensors measuring nitrates and data from the mobile phones of farmers in the region were used. Fang et al. (2014) presented a holistic approach to environmental monitoring and management through an integrated information system that collects data on the regional climate for the city of Xinjiang from various sources including IoT sensors, and related it with ecological response variables such as the primary production and leaf area index. For environmental monitoring, the AirSensEUR project established an affordable open software/hardware multisensor platform, which can monitor air pollution at low concentration levels to create maps of pollution levels in different areas (Kotsev et al. 2016).

A crucial component of any DE system and application is monitoring shifting surface conditions such as erosion on sandy beaches. Pozzebon et al. (2018) presented an Arduino-based system to measure the height of sandy beaches and dunes in real-time. The sensor network uses the ZigBee standard to transmit data, with a GPRS transmitter for sending sensor readings to a MySQL database. Another example is the monitoring of landslides in mountainous areas. Benoit et al. (2015) tested a successful cheap wireless sensor network using XBee for communication and GPS for localization. A thematically related case study is the use of small and inexpensive sensors for monitoring and early-warning systems for floods caused by melting snow in the Quergou River basin (China), as reported by Fang et al. (2015). In addition, changing climate conditions make reliable and efficient management of storm water surges in urban areas important. Rettig et al. (2016) designed and tested a geospatial sensor network for this task, built using common, off-the-shelf components.

With respect to the provision and reception of cultural heritage and cultural services, Chianese et al. (2017) proposed and tested a system that combines business intelligence, Big Data, and IoT data collection to analyze visitor interests and behaviors in a museum. Although IoT devices were only part of the approach, measuring visitor proximity to artworks, their integrated use with other technologies and platforms showcases the strength of a multisensory DE approach.

11.5 Frictions and Synergies Between the IoT and DE

Based on the current technological substrate that provides the initial steps to establish connections between the IoT and DE according to the three cognitive functions (Sect. 11.3), and the presentation of selected case studies (Sect. 11.4), in this section, we (i) carry out a speculative exercise to discuss the main existing limitations and frictions that prevent the IoT and DE from working closer together and (ii) suggest future ways to establish effective communication channels between the two infrastructures.

Before going into detail, it is necessary to establish a fundamental assumption that influences any discussion related to the frictions and synergies between the IoT and DE: the diverging speeds of development of DE and the IoT. New technology and disruptive breakthroughs generally challenge the status quo in any sector, and adopting such improvements can enable more rapid developments and new applications. However, the rapid growth of the IoT field has produced a vast variety of IoT devices and protocols and, consequently, the landscape of IoT-related standards, protocols and specifications is fragmented. For example, a large portion of 'Things' were not originally designed to connect to the Internet; they were later adapted to establish Internet connections by adding connectivity chips via microcontrollers (e.g., Arduino, Raspberry Pi) or through tags (QR Code or RFID). As a result, many different ways to connect hardware and software to enable Internet connectivity were developed and established with no clearly agreed upon consensus and consequently resulted in a lack of interoperability. This example illustrates the great variety and complexity of the IoT universe, where the exponential growth of the IoT is due to the rapid decrease in the size, cost, and energy requirements of sensors, and the ubiquity of network coverage for wireless Internet connections, leading to many standardization efforts following diverging paths. In addition, DE has been traditionally characterized by a slow adaptation of new improvements (López 2011), and thus, the recent technological developments **have not evolved at the same speed in DE as in the IoT**. Noting this fundamental friction, we identify other potential frictions and synergies, which may be considered two sides of one coin, and organize the discussion according to the cognitive functions defined in Sect. 11.3.

11.5.1 Discoverability, Acquisition and Communication of Spatial Information

A direct result of the fragmented standardization context noted above is the **absence of well-accepted global protocols for the discovery of Things**, which also occurs to some extent in DE. Search and discovery is crucial for geo-locating nearby, local, and/or relevant real-world devices and services, a vital step in exploiting sensor data and services to create more advanced knowledge. Early efforts in this direction are discussed in Sect. 11.3.1, but we are still far from a complete solution to this difficult problem, which must be addressed along with the challenges of better description of devices and services and the semantics of the data involved, especially from a geospatial point of view.

Therefore, it remains an open issue to build an IoT-DE ecosystem in a way that will be compatible with standardized IoT reference models and architectures to enable the discovery of relevant sensors (or Things) and related services. Although there are many different scenarios and solutions, several common features can be extracted to find synergies between both infrastructures: the modularity and interoperability of IoT components, open models and architectures, flexible service compositions, integrated security solutions, and semantic data integration. There is an intensified effort regarding the development of architectural frameworks and solutions such as the IEEE or ITU-T models, as well as other related works and approaches developed under the auspices of IETF, W3C, or OASIS. From a DE point of view, associated services for sensor devices and instances are the cornerstone to enable seamless communication and interoperability between the IoT and DE. There are different options such as the SWE and SensorThings API, the latter of which is especially relevant for the establishment of potential solid bridges between the IoT and DE concerning common data models for better data acquisition and unified interfaces for enhanced sensor and service discovery. Some research works have already made substantial progress. Jara et al. (2014) presented a comprehensive framework and architecture to enable discovery over a wide range of technologies and protocols, including legacy systems, and Wang et al. (2015) implemented annotations with an ontology-based semantic service model, SPARQL queries, and geographic indexing to enable sensor discovery in an experimental study, which delivered faster and more accurate responses than other tested approaches.

11.5.2 Spatial Understanding of Objects and Their Relationships

A friction between DE and the IoT is related to the way geographical features are modeled. Traditional GIS data models conceptually abstract the real-world objects into core geometric elements such as points, lines, polygons, and volumes, implemented as raster data models, vector data models, or a combination. These data models were designed to perform spatial analyses such as distance computations and topological operations. Despite these great achievements, GIS (and DE) data models were not designed to cope with the richness and complexity of the interactions between the physical, natural, and social actors that naturally occur in the environment in the way that the IoT potentially can. As noted above, smart devices and Things can 'sense' the environment in a way that was unimaginable before, and, consequently, the streams of rich and finer data acquired by **IoT devices do not fit well with the "coarse-grained" vector/raster data models** widely used in DE applications and systems, as these spatial structures were not intended to handle data with such a high spatio-temporal resolution.

The lack of suitable data models to efficiently manage data at high spatio-temporal resolution highlights **the need for new tools to process data coming from Things and smart devices** in which the modeling of geospatial features has not yet been fully resolved. Moreover, real-time data is often a defining feature in the IoT, as IoT devices and Things can produce data at a high frequency (e.g., data streams), which requires methods for real-time analysis. Therefore, the lack of new algorithms and implementations for real-time computation and processing streams of spatially referenced data sets is a clear limitation. Although some tools can run geospatial queries of stored data, they do not offer ways to analyze data from IoT devices and sensor nodes in real-time (Nittel 2015).

Unlike the IoT, any changes in the DE arena have been more gradual and less frenetic. However, some notable changes indicate the way forward to consolidate potential bridges between DE and the IoT in the midterm and long term. For example, in a Digital Earth Nervous System (De Longueville et al. 2010), **Things could perform basic geospatial operations on sub-networks of Things**, providing processed information for the higher-level elements of a DE. Geometric measurements and basic geospatial analysis are application areas in which Things have been used more widely in recent years (Kamilaris and Ostermann 2018). Similarly, an often overlooked component of IoT applications are the gateway nodes that connect the sensor devices to the wider network. In addition to a simple routing function, these gateways can perform other tasks including exploratory analysis (clustering, event-detection) of incoming data. Rahmani et al. (2018) examined the use of smart gateways in an e-health system that monitors several individual physiological parameters, demonstrating the potential benefits of (spatial) analysis executed directly on smart gateways in the context of DE-related applications such as precision agriculture, environmental monitoring, and disaster management.

The status quo of services for spatial analysis and geoprocessing on the Web is mainly driven by the WPS standard specification (Sect. 11.2.4). However, Herle and Blankenbach (2018) argued that the current WPS standard is not well suited to handle the large amounts of real-time streaming data expected from massive IoT sensor networks. Building on previous work, they extended the WPS with the GeoPipes concept using the GeoMQTT protocol for communication, implementing several smaller proofs-of-concept for application cases such as inverse distance weighting with a sliding window and trajectory data mining. In addition, Armstrong et al. (2018) presented an IoT + CyberGIS system to detect radiation risk and propose that new approaches are needed to integrate the IoT and geospatial analysis and support the fourth scientific paradigm of data-intensive discovery (Hey et al. 2009).

11.5.3 Taking Informed Actions and Acting Over the Environment

In the initial stages of DE, it was thought that sensors could only capture what is happening in the physical environment, i.e., sensors as mere data loggers. The data collected by these sensors are transferred from bottom to top until reaching the SDI repositories. In this sense, the IoT is much more complex because, in addition the feature of acting on the physical environment, the IoT supports communication between devices in the same layer (edge) and complex strategies to determine solutions to real, large problems can be developed. As mentioned above, DE should be adapted to the possibilities that the IoT devices can offer to enrich the capabilities of the current SDIs.

The previously noted heterogeneity problem of connecting IoT devices implies different hardware specifications across the multiple IoT devices. This variety of hardware means that the abovementioned standards cannot work at a low level. This is why the standards mainly define web service interfaces, and connectors or adapters (hub approach) are required to control IoT nodes. Similar to the hub approach, the Sensor Interface Descriptor (SID) solution is a declarative model based on the Sensor ML standard for describing device capabilities (Broering and Below 2010), sensor metadata, sensor commands, and device protocols. In terms of the tasking capability, the SID describes device protocols with the Open Systems Interconnection (OSI) model using an XML schema and thus understanding and adapting the SID may be costly for IoT device manufacturers.

An opportunity that DE can offer the IoT is a global vision on the in situ data that the IoT collects, with the aim of establishing strategies to perform actions in a coordinated manner among the IoT nodes, taking advantage of the ability to act. To conclude, the following Table 11.1 summarizes the frictions and synergies between the IoT and DE.

11.6 Conclusion and Outlook for the Future of the IoT in Support of DE

The concept of combining sensors organized in networks to monitor the environment has been around for decades, and DE has contributed to its expansion. The confluence of new technologies has created a new reality that offers millions of new possibilities, led by the IoT revolution that promises to create a newly interconnected “smart” world (or Earth). After the massive deployment of a ubiquitous array of IoT devices and the

Table 11.1 Detected frictions and synergies between the IoT and DE

	Discoverability, acquisition and communication of spatial information	Understanding spatial objects and their relationships	Taking informed actions and acting over the environment
Frictions	– Absence of well-accepted global protocols for the discovery of Things	– IoT devices do not fit well with coarser vector/raster data models – Lack of tools to process data from Things	– DE has traditionally considered sensors as collectors, with data flowing from bottom to top. – GIS standards must be adapted for each hardware specification
Synergies	– Different standardized IoT models and architectures such as SWE and SensorThings API	– Things can perform basic geospatial operations – Some initiatives have adapted GIS processing standards to support IoT data	– DE provides a global view to establish IoT node strategies to act

impact it made, the world cannot give up being 'online'. Today, the IoT has enabled millions of relationships between objects and Things, so that objects, people, and their environment are more tightly intertwined than ever. Despite the great advances achieved in recent years, like all disruptive innovations, the IoT presents a series of challenges that should be treated as a priority in the coming years, especially in the areas of security, interoperability and standards, privacy, and legal issues. DE can also play a crucial role in handling some of these challenges.

The IoT and DE dichotomy presents various challenges that should be addressed in the near future to create a more beneficial union for both parties: The first challenge is to activate mechanisms to streamline the adaptation of new IoT functionalities from DE. Traditionally, DE is characterized by its comparative inertia to adopt new approaches that imply improvements in terms of performance or usability. Examples include the slow adoption of more flexible interfaces such as the RESTful web interface or data formats that are more suitable for exchange such as JSON in sensor standards such as the SOS specification (Tamayo et al. 2011). The tradeoffs between standardization and disruptive innovation in DE should be carefully discussed by all involved actors to fuel rapid, innovative developments in DE like those in the IoT field. Although the standardization process is key to establishing permanent links between the two infrastructures, it should not slow down innovative changes and technical developments, and standards should be seen as a means to filter out and embrace changes that prove to be useful, effective and valuable for improvement of the IoT-DE ecosystem.

When a technological field grows exponentially, it often leads to heterogeneity and variety in the short term. Within the IoT, this is partly due to the impact that the continuous development and improvement of hardware technology has on IoT devices. Therefore, another challenge to be addressed is the heterogeneity of IoT devices. Although the OGC specifications have helped in the service connection and data/service access levels, the IoT still presents a wide variety of different hardware developments and implementations, most of which are disconnected from the DE infrastructure, and therefore remain invisible for DE applications. The development of ad hoc adapters is one way, at least until a standards consensus is reached in the IoT field, to allow for interaction with the variety of hardware specifications of IoT devices and Things and foster connections between the two infrastructures. This is not an optimal solution since the integration of IoT devices is a challenging and difficult task, but it helps discern the connections and adaptors that may eventually become candidates for standardization bodies.

Throughout this chapter, we revisited many tools that are capable of analyzing spatially referenced data collected by IoT devices. However, the quantity and quality of tools that handle the temporal dimension of data in real time far exceeds those that deal with the spatial dimension. An additional barrier is the large-scale variance in the data models between IoT devices and the decision-making systems that are typically established in DE. Optimal spatial models to handle scale variations can be useful to analyze the information received from IoT devices and obtain a more high-level vision that can be interpreted by decision makers and policy makers. Therefore, investment in the research and development of better tools to spatially analyze IoT data in real time on the edge, fog and cloud scales is a priority in the IoT-DE ecosystem roadmap.

References

ArcGIS Server (n.d.) Retrieved December 10, 2018, from http://enterprise.arcgis.com/es/server/latest/get-started/windows/what-is-arcgis-geoevent-server.htm

Armstrong M P, Wang S, Zhang Z (2018) The Internet of Things and fast data streams: prospects for geospatial data science in emerging information ecosystems. Cartography and Geographic Information Science 1–18. https://doi.org/10.1080/15230406.2018.1503973

Ashton K (2009) That 'internet of things' thing. RFID journal, 22(7): 97–114.

Atzori L, Iera A, Morabito G (2010) The internet of things: a survey. Computer. Networks, 54 (15): 2787–2805.

Ayele E D, Das K, Meratnia N, Havinga P J M (2018) Leveraging BLE and LoRa in IoT network for wildlife monitoring system (WMS), in: 2018 IEEE 4th World Forum on Internet of Things (WF-IoT). Presented at the 2018 IEEE 4th World Forum on Internet of Things (WF-IoT), IEEE, Singapore, pp. 342–348. https://doi.org/10.1109/WF-IoT.2018.8355223

Bakker K, Ritts M (2018) Smart Earth: A meta-review and implications for environmental governance. Global environmental change, 52: 201–211.

Barik R K, Dubey H, Samaddar A B, Gupta R D, Ray P K (2016) FogGIS: Fog Computing for geospatial big data analytics. In Electrical, Computer and Electronics Engineering (UPCON), 2016 IEEE Uttar Pradesh Section International Conference on (pp. 613–618). IEEE.

Bélissent J (2010) Getting clever about smart cities: New opportunities require new business models. Cambridge, Massachusetts, USA.

Benoit L, Briole P, Martin O, Thom C, Malet J P, Ulrich P (2015) Monitoring landslide displacements with the Geocube wireless network of low-cost GPS. Engineering Geology 195, 111–121. https://doi.org/10.1016/j.enggeo.2015.05.020
Botts M, Robin A (2007) OpenGIS sensor model language (SensorML) implementation specification. OpenGIS Implementation Specification OGC, 7(000).
Botts M, Percivall G, Reed C, Davidson J (2008) OGC® sensor web enablement: Overview and high level architecture. In GeoSensor networks (pp. 175–190). Springer, Berlin, Heidelberg.
Brakenridge G R, Anderson E, Nghiem S V, Caquard S, Shabaneh T B (2003) Flood warnings, flood disaster assessments, and flood hazard reduction: The roles of orbital remote sensing.
Broering A, Below S (2010) OpenGIS sensor interface descriptors. Open Geospatial Consortium Inc. 2010-06-30, OGC 10, 134.
Bröring A, Echterhoff J, Jirka S, Simonis I, Everding T, Stasch C, Liang Steve Lemmens R (2011) New generation sensor web enablement. Sensors, 11(3): 2652–2699.
Butler D (2006) Virtual globes: The web-wide world.
Cecchinel C, Jimenez M, Mosser S, Riveill M (2014) An Architecture to Support the Collection of Big Data in the Internet of Things, in: 2014 IEEE World Congress on Services. Presented at the 2014 IEEE World Congress on Services (SERVICES), IEEE, Anchorage, AK, pp. 442–449. https://doi.org/10.1109/SERVICES.2014.83
Chianese A, Marulli F, Piccialli F, Benedusi P, Jung JE (2017) An associative engines based approach supporting collaborative analytics in the Internet of cultural things. Future Generation Computer Systems 66, 187–198. https://doi.org/10.1016/j.future.2016.04.015
Cosgrove-Sacks C (2014) Open Protocols for an Open, Interoperable Internet of Things. Org. for Advancement of Structured Information Standards. www.oasis-open.org/presentations/open-protocols-and-internet-of-things-oasis.ppt (Accessed: 2018-11-10)
Compton M, Barnaghi P, Bermudez L, García-Castro R, Corcho O, Cox S, Huang, V (2012) The SSN ontology of the W3C semantic sensor network incubator group. Web semantics: science, services and agents on the World Wide Web, 17: 25–32.
Cox S (2003) Observation and Measurement (0.9.2 ed., Vol. OGC 03-022r3): OpenGIS Consortium Inc
Díaz L, Granell C, Gould M (2008) Case Study: Geospatial Processing Services for Web-based Hydrological Applications. In: J.T. Sample, K. Shaw, S. Tu, M. Abdelguerfi (Eds). Geospatial Services and Applications for the Internet. New York: Springer Science + Business Media, pp. 31–47
De Smith M, Longley P, Goodchild M (2018) Geospatial analysis – A comprehensive guide (6th ed.). The Winchelsea Press
De Longueville B, Annoni A, Schade S, Ostlaender N, Whitmore C (2010) Digital earth's nervous system for crisis events: real-time sensor web enablement of volunteered geographic information. International Journal of Digital Earth, 3(3): 242–259.
Echterhoff J, Everding T (2008) Opengis sensor event service interface specification. Open Geospatial Consortium Inc., USA, OpenGIS Discussion Paper, OGC, 08-133.
Elahi B M, Romer K, Ostermaier B, Fahrmair M, Kellerer W (2009) Sensor ranking: A primitive for efficient content-based sensor search. In Proceedings of the 2009 international conference on information processing in sensor networks (pp. 217–228). IEEE Computer Society.
Fang S, Xu L, Zhu Y, Liu Y, Liu Z, Pei H, Yan J, Zhang H (2015) An integrated information system for snowmelt flood early-warning based on internet of things. Information Systems Frontiers 17: 321–335. https://doi.org/10.1007/s10796-013-9466-1
Fang S, Xu L, Zhu Y, Ahati J, Pei H, Yan J, Liu Z (2014) An Integrated System for Regional Environmental Monitoring and Management Based on Internet of Things. IEEE Transactions on Industrial Informatics 10: 1596–1605. https://doi.org/10.1109/TII.2014.2302638
Fazio M, Puliafito A (2015) Cloud4sens: a cloud-based architecture for sensor controlling and monitoring. IEEE Communications Magazine 53, 41–47. https://doi.org/10.1109/MCOM.2015.7060517

Gibbons P B, Karp B, Ke Y, Nath S, Seshan S (2003) Irisnet: An architecture for a worldwide sensor web. IEEE pervasive computing, 2(4): 22–33.
Gore A (1998) The digital earth: understanding our planet in the 21st century. Australian surveyor, 43(2): 89–91.
Granell C, Díaz L, Gould M (2010) Service-oriented applications for environmental models: Reusable geospatial services. Environmental Modelling and Software, 25(2): 182–198. https://doi.org/10.1016/j.envsoft.2009.08.005
Granell C (2014) Robust Workflow Systems + Flexible Geoprocessing Services = Geo-enabled Model Web? In: E. Pourabbas (Ed). Geographical Information Systems: Trends and Technologies. Boca Raton: CRC Press (Taylor and Francis Group), pp. 172–204
Gross N (1999) The earth will don an electronic skin. Business Week, August, 30
Grosky W I, Kansal A, Nath S, Liu J, Zhao F (2007) Senseweb: An infrastructure for shared sensing. IEEE multimedia, 14(4).
Grothe M, Broecke J V, Carton L J, Volten H, Kieboom R (2016) Smart Emission-Building a Spatial Data Infrastructure for an Environmental Citizen Sensor Network.
Gubbi J, Buyya R, Marusic S, Palaniswami M (2013) Internet of Things (IoT): A vision, architectural elements, and future directions. Future generation computer systems, 29(7), 1645–1660.
Guinard D, Trifa V, Wilde E (2010) November. A resource oriented architecture for the web of things. In Internet of Things (IOT), 2010 (pp. 1–8). IEEE.
Herle S, Blankenbach J (2018) Enhancing the OGC WPS interface with GeoPipes support for real-time geoprocessing. International Journal of Digital Earth 11, 48–63. https://doi.org/10.1080/17538947.2017.1319976
Havens S (2007) OpenGIS Transducer Markup Language (TML) Implementation Specification, OGC 06-010r6 ed: Open Geospatial Consortium, p. 258.
Hey T, Tansley S, Tolle K (2009) The Fourth Paradigm: Data-Intensive Scientific Discovery. Microsoft Research.
Hofer B, Granell C, Bernard L (2018) Innovation in geoprocessing for a Digital Earth. International Journal of Digital Earth, 11(1): 3–6. https://doi.org/10.1080/17538947.2017.1379154
Intel, n.d. A Guide to the Internet of Things Infographic. Accessed October 29, 2018. Available: https://www.intel.com/content/www/us/en/internet-of-things/infographics/guide-to-iot.html
International Telecommunication Union (2018) "ITU internet reports 2005: the Internet of Things". Accessed October 29, 2018. Available: https://www.itu.int/pub/S-POL-IR.IT-2005/e
IERC (2014) Internet of Things. http://www.internet-of-things-research.eu/about_iot.htm (Accessed: 2018-11-10)
ISO 19156 (2011) Geographic information – Observations and measurements. 2011. https://doi.org/10.13140/2.1.1142.3042.
ITU-T (2012) Recommendation ITU-T Y.2060. Overview of the Internet of Things. ITU-T Study Group 13.
Janowicz K, Haller A, Cox S J, Le Phuoc D, Lefrançois M. (2018) SOSA: A lightweight ontology for sensors, observations, samples, and actuators. Journal of Web Semantics.
Jara A J, Lopez P, Fernandez D, Castillo J F, Zamora MA, Skarmeta A F (2014) Mobile digcovery: discovering and interacting with the world through the Internet of things. Personal and Ubiquitous Computing 18: 323–338. https://doi.org/10.1007/s00779-013-0648-0
Jazayeri M A, Huang C Y, Liang S H (2012) TinySOS: Design and implementation of interoperable and tiny web service for the internet of things. In Proceedings of the First ACM SIGSPATIAL Workshop on Sensor Web Enablement. pp. 39–46. ACM.
Jing C, Wang S, Wang M, Du M, Zhou L, Sun T, Wang J (2018) A Low-Cost Collaborative Location Scheme with GNSS and RFID for the Internet of Things. ISPRS International Journal of Geo-Information 7: 180. https://doi.org/10.3390/ijgi7050180
Jirka S, Bröring A, Stasch C (2009) Discovery mechanisms for the sensor web. Sensors, 9(4): 2661–2681. https://doi.org/10.3390/s90402661

Jo B, Baloch Z (2017) Internet of Things-Based Arduino Intelligent Monitoring and Cluster Analysis of Seasonal Variation in Physicochemical Parameters of Jungnangcheon, an Urban Stream. Water 9: 220. https://doi.org/10.3390/w9030220

Kajimoto K, Davuluru U, Matsukura R, Hund J, Kovatsch M, Nimura K (2017) Web of Things (WoT) architecture. First Public Working Draft, W3C.

Kamilaris A, Ostermann FO (2018) Geospatial Analysis and the Internet of Things. ISPRS International Journal of Geo-Information 7: 269. https://doi.org/10.3390/ijgi7070269

Kamilaris A, Assumpcio A, Blasi A B, Torrellas M, Prenafeta-Boldú F X (2018) Estimating the Environmental Impact of Agriculture by Means of Geospatial and Big Data Analysis: The Case of Catalonia. In From Science to Society (pp. 39–48). Springer, Cham. https://doi.org/10.1007/978-3-319-65687-8_4

Kamilaris A, Yumusak S, Ali M I (2016) WOTS2E: A search engine for a Semantic Web of Things. In Internet of Things (WF-IoT), 2016 IEEE 3rd World Forum on (pp. 436–441). IEEE. https://doi.org/10.1109/wf-iot.2016.7845448

Kamilaris A, Pitsillides A (2014) The Impact of Remote Sensing on the Everyday Lives of Mobile Users in Urban Areas. In Proc. of the 7th International Conference on Mobile Computing and Ubiquitous Networking (ICMU), Singapore, January 2014. https://doi.org/10.1109/ICMU.2014.6799087

Kamilaris A, Papakonstantinou K, Pitsillides A (2014) Exploring the Use of DNS as a Search Engine for the Web of Things. In Internet of Things (WF-IoT), 2014 IEEE World Forum on (pp. 100–105). IEEE. https://doi.org/10.1109/wf-iot.2014.6803128

Kamilaris A, Pitsillides A (2012) Using DNS for Global Discovery of Environmental Services. In WEBIST (pp. 280–284).

Kotsev A, Schleidt K, Liang S, van der Schaaf H, Khalafbeigi T, Grellet S, Lutz M, Jirka S, Beaufils M (2018) Extending INSPIRE to the Internet of Things through SensorThings API. Geosciences, 8(6): 221.

Kotsev A, Schade S, Craglia M, Gerboles M, Spinelle L, Signorini M (2016) Next generation air quality platform: Openness and interoperability for the internet of things. Sensors, 16(3): 403. https://doi.org/10.3390/s16030403

Laska M, Herle S, Klamma R, Blankenbach J (2018) A Scalable Architecture for Real-Time Stream Processing of Spatiotemporal IoT Stream Data—Performance Analysis on the Example of Map Matching. ISPRS International Journal of Geo-Information, 7(7): 238. MDPI AG. http://dx.doi.org/10.3390/ijgi7070238

Lee C K M, Yeung C L, Cheng M N (2015) Research on IoT based Cyber Physical System for Industrial big data Analytics, in: 2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM). Presented at the 2015 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE, Singapore, Singapore, pp. 1855–1859. https://doi.org/10.1109/IEEM.2015.7385969

Lee U, Zhou B, Gerla M, Magistretti E, Bellavista P, Corradi A (2006) Mobeyes: smart mobs for urban monitoring with a vehicular sensor network. IEEE Wireless Communications 13: 52–57. https://doi.org/10.1109/WC-M.2006.250358

Li D, Yao Y, Shao Z, Wang L (2014) From digital earth to smart earth. Chinese Science Bulletin, 59(8): 722–733.

Liang S, Huang C Y, Khalafbeigi T (2016) OGC SensorThings API-Part 1: Sensing. OGC® Implementation Standard.

Liang S, Khalafbeigi T (2018) OGC SensorThings API Part 2: Tasking, Version 1.0, Open Geospatial Consortium. OGC. Candidate standard

López T S (2011) RFID and sensor integration standards: State and future prospects. Computer Standards and Interfaces, 33(3): 207–213.

Lü G, Batty M, Strobl J, Lin H, Zhu A-X, Chen M (2019) Reflections and speculations on the progress in Geographic Information Systems (GIS): a geographic perspective. International Journal of Geographical Information Science 33:2, 346–367, https://doi.org/10.1080/13658816.2018.1533136

McCullough A, Barr S, James P (2011) A Typology of Real-Time Parallel Geoprocessing for the Sensor Web Era. In Integrating sensor web and web-based geoprocessing. CEUR Workshop proceedings (Vol. 712).
Minerva R, Biru A, Rotondi D (2015) Towards a definition of the Internet of Things (IoT). IEEE Internet Initiative, 1, 1–86.
Miz V, Hahanov V (2014) Smart traffic light in terms of the cognitive road traffic management system (CTMS) based on the Internet of Things, in: Proceedings of IEEE East-West Design and Test Symposium (EWDTS 2014). Presented at the 2014 East-West Design and Test Symposium (EWDTS), IEEE, Kiev, Ukraine, pp. 1–5. https://doi.org/10.1109/EWDTS.2014.7027102
Newman D (2017) "The Top 8 IoT Trends For 2018". Forbes. Accessed October 29, 2018. Available: https://www.forbes.com/sites/danielnewman/2017/12/19/the-top-8-iot-trends-for-2018/#1e35e7f467f7
Nittel S (2015) Real-time sensor data streams. SIGSPATIAL Special, 7(2): 22–28.
OGC (2005) OpenGIS Web Processing Service. OpenGIS Discussion Paper. Schut, P. and Whiteside, A. (eds.).
Oksanen J, Bergman C, Sainio J, Westerholm J (2015) Methods for deriving and calibrating privacy-preserving heat maps from mobile sports tracking application data. Journal of Transport Geography, 48: 135–144. https://doi.org/10.1016/j.jtrangeo.2015.09.001
O'Reilly T (2010) "OGC® puck protocol standard version 1.4." Open Geospatial Consortium, OGC Encoding Standard OGC.
OData. "OData – the Best Way to REST". www.odata.org. Accessed 24 Dec. 2018.
Pradilla J, Palau C, Esteve M (2015) SOSLite: lightweight sensor observation service (SOS). IEEE Latin America Transactions, 13(12): 3758–3764.
Pradilla J, González R, Esteve M, Palau C (2016) Sensor Observation Service (SOS)/Constrained Application Protocol (CoAP) proxy design. In Electrotechnical Conference (MELECON), 2016 18th Mediterranean (pp. 1–5). IEEE.
Perez A J, Labrador M A, Barbeau S J (2010) G-sense: a scalable architecture for global sensing and monitoring. IEEE Network, 24(4). https://doi.org/10.1109/MNET.2010.5510920
Portele C, van Genuchten P, Verhelst L, Zahnen A (2016) https://geo4web-testbed.github.io/topic4/
Pfisterer D, Romer K, Bimschas D, Kleine O, Mietz R, Truong C, Hasemann H, Pagel M, Hauswirth M, and Karnstedt, M. 2011. SPITFIRE: toward a semantic web of things. IEEE Communications Magazine, 49(11): 40–48. https://doi.org/10.1109/mcom. 2011.6069708
Pschorr J, Henson C A, Patni H K, Sheth A P (2010) Sensor discovery on linked data. Kno.e.sis, Technical Report.
Puiu D, Barnaghi P, Tönjes R, Kümper D, Ali M I, Mileo A, Alessandra X, Parreira J, Marten K, Gao F (2016) Citypulse: Large scale data analytics framework for smart cities. IEEE Access, 4: 1086–1108. https://doi.org/10.1109/ACCESS.2016.2541999
Pozzebon A, Andreadis A, Bertoni D, Bove C (2018) A Wireless Sensor Network Framework for Real-Time Monitoring of Height and Volume Variations on Sandy Beaches and Dunes. ISPRS International Journal of Geo-Information 7: 141. https://doi.org/10.3390/ijgi7040141
Rahmani A M, Gia T N, Negash B, Anzanpour A, Azimi I, Jiang M, Liljeberg P (2018) Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach. Future Generation Computer Systems 78: 641–658. https://doi.org/10.1016/j.future.2017.02.014
Rettig A J, Khanna S, Beck R A, Wojcik Q, McCane C A (2016) Monitoring permeable paver runoff with an open-innovation geospatial sensor network. International Journal of Digital Earth 9: 30–46. https://doi.org/10.1080/17538947.2014.965880
Rodríguez-Pupo L E, Casteleyn S, Granell C (2017) On Metrics for Location-aware Games. ISPRS International Journal of Geo-Information 6(10): 299.http://dx.doi.org/10.3390/ijgi6100299
Rieke M, Bigagli L, Herle S, Jirka S, Kotsev A, Liebig T, Malewski C, Paschke T, Stasch C (2018) Geospatial IoT—The Need for Event-Driven Architectures in Contemporary Spatial Data Infrastructures. ISPRS International Journal of Geo-Information, 7(10): 385. MDPI AG. Retrieved from http://dx.doi.org/10.3390/ijgi7100385

Song W, Shirazi B, Kedar S, Chien S, Webb F, Tran D, Davis A, Pieri D, LaHusen R., Dzurisin D (2008) Optimized autonomous space in-situ sensor-web for volcano monitoring. In Aerospace Conference, 2008 IEEE (pp. 1–10). IEEE.

Sawant S, Durbha S S, Jagarlapudi A (2017) Interoperable agro-meteorological observation and analysis platform for precision agriculture: A case study in citrus crop water requirement estimation. Computers and Electronics in Agriculture 138: 175–187. https://doi.org/10.1016/j.compag.2017.04.019

Sawant S A, Adinarayana J, Durbha S S (2014) KrishiSense: A semantically aware web enabled wireless sensor network system for precision agriculture applications, in: 2014 IEEE Geoscience and Remote Sensing Symposium. Presented at the IGARSS 2014 – 2014 IEEE International Geoscience and Remote Sensing Symposium, IEEE, Quebec City, QC, pp. 4090–4093. https://doi.org/10.1109/IGARSS.2014.6947385

Schade S, Díaz L, Ostermann F O, Spinsanti L, Luraschi G, Cox S, Nuñez M, De Longueville B (2013) Citizen-based sensing of crisis events: sensor web enablement for volunteered geographic information. Applied Geomatics 5: 3–18. https://doi.org/10.1007/s12518-011-0056-y

Segura Garcia J, Pérez Solano J, Cobos Serrano M, Navarro Camba E, Felici Castell, S, Soriano Asensi A, Montes Suay F (2016) Spatial Statistical Analysis of Urban Noise Data from a WASN Gathered by an IoT System: Application to a Small City. Applied Sciences 6: 380. https://doi.org/10.3390/app6120380

Seydoux N, Drira K, Hernandez N, Monteil T (2016) IoT-O, a core-domain IoT ontology to represent connected devices networks. In European Knowledge Acquisition Workshop, Springer, Cham, pp. 561–576

Shafi M, Molisch A F, Smith P J, Haustein T, Zhu P, Silva P D, Tufvesson F, Benjebbour A, Wunder G (2017) 5G: A tutorial overview of standards, trials, challenges, deployment, and practice. IEEE Journal on Selected Areas in Communications, 35(6): 1201–1221

Sheth A P (2018) Semantic sensor web. https://corescholar.libraries.wright.edu/knoesis/752/

Shi W, Cao J, Zhang Q, Li Y, Xu L (2016) Edge computing: Vision and challenges. IEEE Internet of Things Journal, 3(5): 637–646.

Simonis I (2006) OGC Sensor Alert Service Candidate Implementation Specification, OGC document number 06-028r3, http://portal.opengeospatial.org/files/?artifact_id=15588

Simonis I (2007) Sensor Planning Service Implementation Specification, Version 1.0. OGC document, 7.

Tamayo A, Granell C, Huerta J (2011) Analysing performance of XML data binding solutions for SOS applications. In Proceedings of Workshop on Sensor Web Enablement (Vol. 2011).

Trilles S, Belmonte O, Diaz L, Huerta J (2014) Mobile access to sensor networks by using GIS standards and restful services. IEEE Sensors Journal, 14(12): 4143–4153.

Trilles S, Luján A, Belmonte Ó, Montoliu R, Torres-Sospedra J, Huerta J (2015a) SEnviro: a sensorized platform proposal using open hardware and open standards. Sensors, 15(3): 5555–5582.

Trilles S, Schade S, Belmonte Ó, Huerta J (2015b) Real-time anomaly detection from environmental data streams. In AGILE 2015 (pp. 125–144). Springer, Cham.

Trilles S, Belmonte Ó, Schade S, Huerta J (2017) A domain-independent methodology to analyze IoT data streams in real-time. A proof of concept implementation for anomaly detection from environmental data. International Journal of Digital Earth, 10(1): 103–120.

Tao C, Xiang L (2010) Municipal Solid Waste Recycle Management Information Platform Based on Internet of Things Technology, in: 2010 International Conference on Multimedia Information Networking and Security. Presented at the 2010 International Conference on Multimedia Information Networking and Security, IEEE, Nanjing, China, 2010

Van Setten M, Pokraev S, Koolwaaij J (2004) Context-aware recommendations in the mobile tourist application COMPASS. In International Conference on Adaptive Hypermedia and Adaptive Web-Based Systems (pp. 235–244). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-27780-4_27

Van der Zee E, Scholten H (2014) Spatial Dimensions of Big Data: Application of Geographical Concepts and Spatial Technology to the Internet of Things. In: Bessis N., Dobre C. (eds) Big

Data and Internet of Things: A Roadmap for Smart Environments. Studies in Computational Intelligence, vol 546. Springer, Cham. https://doi.org/10.1109/MINES.2010.155
Wang F, Yuan H (2010) Challenges of the sensor web for disaster management. International Journal of Digital Earth, 3(3): 260–279.
Wang W, De S, Cassar G, Moessner K (2015) An experimental study on geospatial indexing for sensor service discovery. Expert Systems with Applications 42: 3528–3538. https://doi.org/10.1016/j.eswa.2014.11.058
Wang H, Tan C C, Li Q (2008) Snoogle: A search engine for the physical world. In INFOCOM 2008. The 27th Conference on Computer Communications. IEEE (pp. 1382–1390). IEEE. https://doi.org/10.1109/infocom.2008.196
Worboys M F, Duckham M (2004) GIS: A Computing Perspective, Second Edition. Boca Raton: CRC Press.
Xu Z, Zhang H, Sugumaran V, Choo K K R, Mei L, Zhu Y (2016) Participatory sensing-based semantic and spatial analysis of urban emergency events using mobile social media. EURASIP Journal on Wireless Communications and Networking, 2016(1): 44. https://doi.org/10.1186/s13638-016-0553-0
Yue S, Chen M, Wen Y, Lu G (2016) Service-oriented model-encapsulation strategy for sharing and integrating heterogeneous geo-analysis models in an open web environment. ISPRS Journal of Photogrammetry and Remote Sensing, 114: 258–273. https://doi.org/10.1016/j.isprsjprs.2015.11.002
Zhao P, Foerster T, Yue P (2012) The Geoprocessing Web. Computers and Geosciences, 47: 3–12. https://doi.org/10.1016/j.cageo.2012.04.021
Zhou Y, De S, Wang W, Moessner K (2016) Search techniques for the web of things: A taxonomy and survey. Sensors, 16(5): 600. https://doi.org/10.3390/s16050600
Zook M, Graham M, Shelton T, Gorman S (2010) Volunteered geographic information and crowdsourcing disaster relief: a case study of the Haitian earthquake. World Medical and Health Policy, 2(2):7–33. https://doi.org/10.2202/1948-4682.1069

Carlos Granell currently holds a 5 year Ramón y Cajal post-doctoral fellowship at the Universitat Jaume I of Castellón, Spain. His interests lie in multi-disciplinary application of GIS (Science/Systems), spatial analysis & visualization of streams of sensor- and/or user-generated geographic content, reproducibility research practices, and finding synergies between citizen science projects, education and dissemination of science.

Andreas Kamilaris is an Assistant Professor at the Computer Science department of the University of Twente, the Netherlands and the leader of the SuPerWorld research laboratory at the Research Centre on Interactive Media, Smart Systems and Emerging Technologies (RISE), in Nicosia Cyprus. Andreas has published more than 60 research papers, has received several awards and has co-founded three start-ups. His research interests include geospatial analysis, Internet/Web of Things, machine learning and smart ecosystems.

Alexander Kotsev is a scientific project officer at the European Commission's Joint Research Centre. He holds a Ph.D. in Geographic Information Science from the University of Sofia. His research interests include data analysis, sensor web, open source technology and architectures for knowledge management.

Frank O. Ostermann is an Assistant Professor at the faculty ITC at the University of Twente (Netherlands). His research focuses on handling novel sources of geoinformation from citizens and small in-situ sensors. He received his Ph.D. from the University of Zurich (Switzerland), and worked as a post-doctoral researcher at the European Commission's Joint Research Center.

Sergio Trilles received his Ph.D. in Integration of Geospatial Information from the Jaume I University in 2015. He had the opportunity to work four months as researcher in the Digital Earth and Reference Data Unit of the European Commission's Joint Research Centre (JRC). Currently he is a postdoc researcher at the GEOTEC group.

Chapter 12
Social Media and Social Awareness

Xinyue Ye, Bo Zhao, Thien Huu Nguyen and Shaohua Wang

Abstract The human behaviors and interactions on social media have maintained themselves as highly dynamic real-time social systems representing individual social awareness at fine spatial, temporal, and digital resolutions. In this chapter, we introduce the opportunities and challenges that human dynamics-centered social media bring to Digital Earth. We review the information diffusion of social media, the multi-faced implications of social media, and some real-world cases. Social media, on one hand, has facilitated the prediction of human dynamics in a wide spectrum of aspects, including public health, emergency response, decision making, and social equity promotion, and will also bring unintended challenges for Digital Earth, such as rumors and location spoofing on the other. Considering the multifaceted implications, this chapter calls for GIScientists to raise their awareness of the complex impacts of social media, to model the geographies of social media, and to understand ourselves as a unique species living both on the Earth and in Digital Earth.

Keywords Social media · Human dynamics · Social awareness · Location spoofing

12.1 Introduction: Electronic Footprints on Digital Earth

Geo-positioning system-enabled instruments can record and reveal personal awareness at fine spatial, temporal, and digital resolutions (Siła-Nowicka et al. 2016; Li et al. 2017; Ye and Liu 2019). With an exponential growth, human dynamics data are retrieved from location-aware devices, leading to a revolutionary research agenda regarding what happens where and when in the everyday lives of people in both real and virtual worlds (Batty 2013; Yao et al. 2019). Many location-based social media

X. Ye (✉) · S. Wang
College of Computing, New Jersey Institute of Technology, Newark, USA
e-mail: xye@njit.edu

B. Zhao
Department of Geography, University of Washington, Seattle, USA

T. H. Nguyen
Department of Computer and Information Science, University of Oregon, Eugene, USA

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_12

(LBSM) instances have been gaining popularity, fostering the emergence of fine-grained georeferenced social media content through these personalized devices (Liu et al. 2018a, b). The proliferation of LBSM enables researchers and practitioners to efficiently track a large and growing number of human action and interaction records over time and space to develop insights and enhance decision-making process from individual to global levels. The patterns and trends produced by LBSM can identify the movement of active social media users and aid in inferring demographics and related infrastructures. The collected data on users' physical and virtual activities facilitate the in-depth understanding of human dynamics from various aspects (Barabasi 2005; Shaw et al. 2016). The large volumes of such user-generated locational and contextual information are especially beneficial to studies relevant to the evolution of population size and human settlement structure as well as highly topical subjects such as traffic and epidemiological forecasting. For instance, real-time customer shopping behaviors might be rapidly identified by searching specific keywords in tweets, which allows for urban researchers and business analysts to monitor the fine-scale dynamics of economic geography and market outcome (Ye and He 2016). This new data landscape might not directly provide an ultimate solution to long-standing social or economic issues, but can increasingly shed light on many societal characteristics that are otherwise difficult to discover using traditional questionnaires or surveys.

Human actions and interactions in the digital form as well as frequent status updates can manifest themselves as highly dynamic real-time social systems, which enable the government to formulate appropriate policies for the relevant groups and targeted communities (Shi et al. 2018; Wang and Ye 2018). The electronic footprints and perceptions left by social media users and derivatives of complicated social networks can be utilized to enhance the design of location-based services (Ye and Lee 2016). Hence, the increasing demands in mapping and analyzing social media data call for more innovative conceptual and technological advances in visual and computational methods. These research challenges and opportunities can facilitate a paradigm shift in the broader social science disciplines in this new form of data landscape. Social media messages can depict the interconnected patterns and relationships between cyberspace and physical space, and can also be distributed instantly to a large number of users globally, who may belong to different virtual communities (Shelton et al. 2015).

Geographic information has traditionally been spread by governments or industries in a top-down manner; but its broadcast is much faster through social media than official agencies. The dramatic transition towards bottom-up digital dissemination has challenged these official or professional processes. Individuals can utilize the power of volunteered geographic information to minimize the difference and/or quality between experts and nonexperts in the context of generating a large collection of user-described features and numerous georeferenced citizen observations on socio-economic phenomena. With social media platforms becoming increasingly location-enabled, users can share geo-tagged information about their own lives and, as a result, rich content about large populations can be aggregated for social and behavioral studies (Sui and Goodchild 2011). Such a practice facilitates the policy transition from long-term to short-term action with a new perspective of understanding, visualizing, and analyzing human dynamics (Batty 2013).

The use of LBSM content represents a significant methodological advancement in social sciences and humanities research, providing rich content regarding human-environment interaction with locational estimation in ubiquitous/pervasive computing. It can efficiently assist place-based policy interventions in a timely fashion. Prompt and rigorous detection of emerging social and economic events calls for more robust algorithms to support such unprecedented research efforts in both qualitative and quantitative analyses. However, challenges and difficulties remain in processing user-generated messages to derive effective and high-quality information, considering the complex syntax and context embedded in social media messages. Additionally, if data analytics cannot be effectively conducted, the expected results could lose value for decision makers. These issues must be addressed to realize the potential of social media analytics.

Considering the above-mentioned issues, the remainder of this chapter is organized as follows. Section 12.2 describes the multifaceted implications of social media. Regarding social media, the unprecedented opportunities to predict human dynamics are introduced in Sect. 12.3, while multiple challenges are listed in Sect. 12.4. Then, the implications of these opportunities and challenges are further discussed in Sect. 12.5, followed by a conclusion in Sect. 12.6.

12.2 Multifaceted Implications of Social Media

Value systems are fundamental to anything we do. Today, the rapid proliferation of social media has greatly affected us and almost every aspect of human society. Confronted with this complicated and unstoppable interaction, we employ value structures to holistically discover the implications of social media, especially the unintended but vital ones. McLuhan's (1975) law of media is frequently utilized to capture the social consequences of various media. Tuan (2003) also proposed the psychology of power to unveil the internal logic of human's perceptions of places, and Ihde (1990) contemplated how technology mediates between human beings and the world from a phenomenological perspective.

Among these value structures, we employ Ihde's amplification-reduction structure to investigate the opportunities and challenges brought by social media. This structure reveals how technology (including social media) amplify and simultaneously reduce a certain human experience. The amplified and reduced experiences are intertwined and interrelated. More significantly, the amplified human experience is obvious whereas the reduced human experience is undiscoverable and easily ignored. Though Ihde only suggested applying this structure to the human experience, it can also be applied to understand the social implications of the investigating object. Through this structure, the opportunities and challenges of the social implications can be revealed. For example, during the 2008 Olympic Games, social media was touted as a tool of freedom to enable the general public to express their concerns about the air pollution issues in major Chinese cities. If we acknowledge the promotion of free speech as the opportunities brought about by social media, the hidden challenges can be revealed through this value structure—social media can also be used as a tool of surveillance by big brother to control the discussion on air pollution as well as a medium of advertisement by private companies to sell relevant products (e.g., masks) to prevent air pollution-related symptoms. The implications of social media are multifaceted. Therefore, the value structure can be applied to examine the impacts of social media on the rapidly evolving Digital Earth. In the following sections, we discuss the opportunities provided by social media as well as the potential challenges.

12.3 Opportunities: Human Dynamics Prediction

As a newly chartered territory for human activities, social media has resulted in tremendous electronic footprints. Such footprints represent a large number of the population and can be used to predict human dynamics on the ground via the relationships between the spread of information, user characteristics, and message contents. In this section, we discuss how social media can be used for different aspects of human society, including public health, emergency response, decision making, and social equity promotion.

12.3.1 Public Health

Social media platforms can be used to mitigate the spread of pandemics and associated anxiety. Scholars have used sentiment analysis and spatial analysis to examine how social media communication conveys information about contagious and infectious diseases and alerts the public, through identifying, tracking, and visualizing the behavioral patterns of users (Zadeh et al. 2019). For instance, Ye et al. (2018a, b) explored public health-related rumors during disease outbreaks and evaluate how such media framing sets the tone negatively, affecting the quality of disease outbreak detection and prediction, using the diffusion of Ebola rumors in social media networks as a case study. Sharma et al. (2017) find that the inaccurate Facebook posts are more popular than those with accurate and relevant information about the Zika virus. Villar and Marsh (2018) studied the impact of social media health communication of Ebola and Zika, concluding that the effect relies on users' attitudes and trust towards authorities and the media. Average citizens and ordinary social media users have very limited knowledge regarding the accuracy and relevancy of infectious diseases spreading over time and across space as well as concerning complications. As a force in health communication, social media data could be utilized to define a temporal extent of the infection and to populate a spatial database of reported occurrences of the disease. Additionally, social media data can be used to track and predict the emergence and spread of infectious diseases and distribution across various spatial and temporal scales. As a self-reported volunteered information platform and useful surveillance tool, social media feeds outperform those from official or government outlets in timeliness. They can also aid in gaining insights into the opinions and perceptions of the public.

12.3.2 Emergency Response

The use of massive computer-mediated communication in emergency response and disaster management has captured considerable interest from both the general public and decision makers. Social media enables fast interpersonal communication during crises through information dissemination, early warnings, environmental awareness, and public participation in disaster-affected areas, allowing for emergency workers to respond more speedily and capably (Hashimoto and Ohama 2014; Finch et al. 2016). As Yin et al. (2012) argue, "this growing use of social media during crises offers new information sources from which the right authorities can enhance emergency situation awareness. Survivors in the impacted areas can report on-the-ground information about what they are seeing, hearing, and experiencing during natural disasters. People from surrounding areas can provide nearly real-time observations about disaster scenes, such as aerial images and photos." Moreover, since social media users can access information posted by official agencies through following their accounts, organizations and agencies can leverage social media as a platform to post authoritative situational announcements and communicate with the public in emergency situations and to potentially retrieve and verify on-the-ground information using the public as the information source (Wang et al. 2016). Palen et al. (2009) examined the consequences of digital communication and information sharing on emergency response in the context of the Virginia Tech massacre. Chen et al. (2016) proposed real-time geo-tagged tweet collection and recording in a distributed geodatabase as well as real-time data redistribution using a Web GIS application. This system was applied to a hypothetical mass evacuation using tweets from Hurricane Joaquin in 2015.

12.3.3 Decision Making

As a new kind of user-generated geospatial information, social media data could be invaluable to political agenda-setting that needs to be aware of location-based topic distribution. For example, the data could help political strategists analyze the tweets of residents or voters in a given geographical area. Politicians can gauge people's reactions by monitoring the communication among Twitter accounts regarding policy issues. Ye et al. (2017) employed voting tweets regarding a water bond in California to highlight place-based situational awareness. Convention and visitors' bureaus may focus on 'hot button' issues in certain places within their cities or regions. These data could provide operational indicators about places that are most visited or preferred by visitors, which can inform the marketing strategies relevant to these locations. Local governments could analyze social media messages to determine whether a proposed construction project would be favored by the public or if other proposed projects would be perceived positively by their constituents. Ye et al. (2018a, b) examined how the Multilevel Model of Meme Diffusion (M3D) captures the debates regarding death penalty abolishment across space. At the intracity scale, Liu et al. (2018a, b) assessed the utility efficiency of subway stations in a Chinese city by matching the capacity of train services and the travel needs using social media data. Deng et al. (2018) analyzed how geotagged tweets are associated with hourly electric consumption at the building level, given the assumption that tweeting behavior is highly related to human activities.

12.3.4 Social Equity Promotion

Most social media platforms such as Twitter, Facebook, or Instagram are designed for the general public; few are dedicated to specific groups (e.g., LGBTQ, photographers, natural disaster victims, etc.). An in-depth analysis and visualization of the specific groups can promote social equity among different groups. Social awareness of where they are is the first and foremost step in enabling local residents and governments to recognize the necessity to treat these underrepresented populations equally. For example, Jack'd, a dedicated gay social networking app, enables its users to communicate online with those who are physically nearby. Through collecting online locational information from Jack'd, a 3D distribution of the gay community in Beijing were visualized (Zhao et al. 2017). By overlapping this distribution with landmarks such as major roads, university campuses, shopping malls and gay-friendly places (e.g., gay bars, gay saunas, gay-friendly gyms, gay-friendly parks and public restrooms attracting gay activities, etc.), the characteristics of this underrepresented group's distribution can be revealed. Gay people in Beijing primarily concentrate in the northwestern and eastern parts of the city. The northwestern area is the center for higher education, with several famous universities. In the eastern area of Beijing, the area from Sanlitun to Worker's stadium is acknowledged as a recreation center for LGBT people. To the south, a few famous gay-friendly residential communities are surrounded by gay saunas; to the east, there are several high-end residential communities and shopping malls in the Guomao and Sihui subdistricts. This 3D distribution reveals a hot spot of gay activities in the Tongzhou district. This may result from the relatively low house rent and convenient accessibility to Chaoyang and other local urban centers for hangouts. Through this 3D distribution of the gay community's electronic footprint, the local public health agencies can provide corresponding services for the gay community and organize more targeted activities as an effective means to promote social equity.

12.4 Challenges: Fake Electronic Footprints

In addition to those obvious opportunities in human dynamics prediction, challenges inherently in social media are often ignored. As Chun et al. (2019) argue, "uncertainty and context pose fundamental challenges in GIScience and geographic research. Geospatial data are imbued with errors (e.g., measurement and sampling) and various types of uncertainty that often obfuscate any understanding of the effects of contextual or environmental influences on human behaviors and experiences." Although social media has been touted as a platform to authentically present human trajectories and their mobilities, rumors, spoofings and privacy concerns, not limited to the physical world, are also exist on Digital Earth. In this sense, We cannot immediately treat social media messages as accurate and credible without considering the above-mentioned issues.

12.4.1 Rumors

The unmoderated nature of social media user's posting behavior might lead to the accumulation of invalidated and unverified information and news involving speculation and uncertainty regarding social events (Ye et al. 2018a, b). Jones et al. (2017) found those who relying on social media for updates of a campus lockdown tend to suffer from greater distress due to their increased exposure to conflicting content in social media channels. Rumors are considered messages or forms of interaction among people about certain events that may not be true. As a nonprofessional medium, social media platforms can spread rumors. However, some information from reliable sources can minimize rumor propagation, lowering the level of anxiety in the virtual community. Zubiaga et al. (2018) noted that the openness of social media platforms also enables the study of user behavior on sharing and discussing both long-standing and newly emerging rumors based on natural language processing and data mining methods, especially for four components: rumor detection, rumor tracking, rumor stance classification, and rumor veracity classification.

12.4.2 Location Spoofing

Location spoofing is a deliberate geographic practice to disguise one's actual location with inconsistent locational information (Zhao and Sui 2017). It facilitates the spoofer to virtually travel to places of interest for various purposes. For smartphones, the spoofing mechanism can be divided into three steps, (1) blocking the positioning service of a smartphone to acquire the actual locational information, (2) generating inconsistent locational information, and (3) transmitting it to an operating LBSM app (e.g., Twitter, Facebook, Pokémon Go, etc.) on a smartphone (Zhao and Zhang

2018). As a result, the LBSM app mistakes the fake location as where the operating smartphone really is. Specifically, the positioning service relies on a hybrid approach that integrates three major positioning techniques: built-in GPS, surrounding WiFi network triangulation, and cellular tower network triangulation. For these three techniques, the more accurate, the higher priority in deciding the final result. In practice, the fundamental function of location spoofing is to downgrade the accuracy of the positioning technique or totally block the positioning function. There are three common location spoofing techniques, in terms of falsifying the MAC addresses of surrounding WiFi routers, spoofing GPS signals in the environment, and mocking in-transit locational information. The last method is predominantly adopted by dedicated mobile android apps for location spoofing. Such apps enable users to virtually visit a place other than the actual location. An example is presented below to clarify this.

In reaction to the 2009 presidential election in Iran, the government of Iran regularly monitors all activities on social networks (Ansari 2012). During the campaign, social network sites were suddenly blocked, and online political activity became the target of harsh criticism and reprisals from the government. To prevent this surveillance and protect online protestors, many internationally based Green Movement supporters spread disinformation over Twitter to mislead local police. Foreign supporters who were not in Iran decided to set their online locations to Tehran to protect those who were tweeting from Tehran. This strategy may have helped some Iranian opposition leaders avoid persecution, but also made it impossible to understand the real impacts of Twitter on the protest.

12.4.3 Privacy Abuse

When users share content and their data on social media, there is a risk that such content and data are collected and exploited in a way that is not expected by the users. This poses a serious challenge in terms of privacy for user data and calls for the responsibility of the network administrators, researchers and users to preserve privacy in social networks. Two broad classes of privacy issues in social networks—user-user privacy and user-third party privacy—are discussed below.

In social media, one user might share content about another user or party. Although this mechanism helps spread the content over the networks efficiently, it inherently presents a tremendous risk for privacy violation. For instance, your friends might share a picture you posted, showing you were in a restaurant with another friend. The picture sharing might be done without your consent and accidentally reveal your location, private information that you do not want to share beyond your friend list. To prevent such privacy breaches, social media administrators have implemented mechanisms for users to make complaints and request that the content be removed from the networks. However, before the content can be reviewed and revoked, it might have caused some detrimental consequences for the users. It would be more effective if such content dissemination was validated at the very beginning. Addressing the user-user privacy issue requires collaboration among scientists from different disciplines, including computer scientists, GIScientists, and psychologists. For example, Kekulluoglu et al. (2017) studied a hybrid negotiation architecture with a reciprocity mechanism to mimic the social responsibility in reality, and a credit system was used to encourage agents/users to respect other's privacy in social media.

Regarding user-third party privacy, content and data generated by social media users might be collected by different third parties for various purposes, potentially causing serious data leaks and violating the privacy of users. A retailer might retrieve user profiles and posts to deliver appropriate ads to the users or an upstart voter-profiling company could exploit such information to characterize the personalities of users and influence their voting decisions (e.g., the recent Facebook privacy crisis and data leak with Cambridge Analytica on American elections described in Rosenberg et al. (2018)). Another example is researchers who query user data to infer various user characteristics (e.g., depression, drug abuse) (Choudhury et al. 2013). While such inferences can provide valuable insights into different social problems and support monitoring systems for social issues, the leaks of such inferred information for specific users can cause biases and affect the users' ability to participate in social activities (e.g., jobs, school admission). Consequently, it is important to develop technological strategies to ensure privacy in user data-related activities in social media. The Future of Privacy Forum and DataGuidance (2018) delivered the report "Comparing privacy laws: GDPR v. CCPA." This report compares the European Union's General Data Protection Regulation (GDPR) effective on May 25, 2018, and the California Consumer Privacy Act of 2018 (CCPA) scheduled to be in effect on January 1, 2020. Both laws would also fundamentally influence social media platforms in collecting/sharing/employing users' data online and offline.

12.5 From Awareness to Action

A close scrutiny of the opportunities and challenges would raise our awareness of the potential capacity of social media in understanding human dynamics. As Yang et al. (2016, p. 61) argued, "the convergence of social media and GIS provides an opportunity to reconcile space-based GIS and place-based social media." Driven by this awareness, GIScientists should take actions to model the geographies of social media, propose innovative approaches to location spoofing screening and connect the virtual world in social media with the real word to better explain social media phenomena.

12.5.1 Modeling the Geographies of Social Media

Tracking and predicting the diffusion of social media information from a neighborhood to a global scale raises a series of questions such as where and when certain topics will be discussed and become popular. Sui and Goodchild (2011) suggested two hypotheses to test the nature of social media message diffusion such as geo-tagged hashtags spread through Twitter. The spatial influence model states that the spatially nearby locations tend to be impacted in the near future, and the community affinity influence model asserts that such dissemination would occur between functionally connected places. However, the reality is usually a combination of these two models. Such predictions will be useful for policymakers to estimate the spatial and functional influence of economic downturns facilitated by supply-chain networks. The community affinities are expected to enhance the prediction power of purely spatial models.

12.5.2 Detecting Location Spoofing Through Geographic Knowledge

If we examine location spoofing from the traditional standard of scientific data, it is highly unlikely that such "fake" information is generated by environmental uncertainties, measurement uncertainties, or limited knowledge about measurement (Zhang and Goodchild 2002). Today, location spoofing cannot simply be treated as fake data, as these data are associated with complicated generative motivations from different stakeholders, governments, local business or average social media users. To identify location spoofing, it is necessary to determine the motivations why the author produces that location, and then judge whether it is spoofed or not.

Therefore, we must seek appropriate solutions to the positioning inconsistency and the motivations for spoofing. Usually, self-reporting (e.g., survey, questionnaire) or observations can qualitatively collect and interpret human motivations that trigger the generation of positional inconsistency. However, in practice, it is difficult to measure the real motivation: admittedly, the survey or questionnaire participants might not report their true intentions of location spoofing due to the fear of being recognized as location spoofers or rumormongers.

The positioning inconsistency in spoofing can be quantitatively detected. Theoretically, any spoofing detection is supposed to unveil a certain underlying positioning inconsistency. As Goodchild (2013) indicated, a geospatial accuracy model interprets how a world is constructed geographically. In this sense, spoofing detection is meant to detect scenarios that do not follow the way in which the world is geographically constructed. One crucial theoretical framework to build up the geographic truth is Hägerstrand's Time Geography (1970). This analytical framework conceptualizes the trajectory of each individual as a life path, which is restricted by several predefined human behavioral constraints in space and time. Meanwhile, a series of analytical tools to measure human dynamics are provided by Time Geography, such as space-time path, prism, and cube. Zhao and Sui (2017) provided a Bayesian time geographic estimation approach to determine the places that an examined user is unlikely to appear. Time-geographic density estimation (TGDE) was used to model the human appearance in a region over time. TGDE can convert trajectories (e.g., a time sequence of historical geo-tags from an individual) to a visiting probability distribution of spatial positions over time. This model can effectively convey the behavioral constraints and describe where and when an individual is more likely to visit. A location with a lower probability value is more likely to be spoofed. Moreover, with the rise of deep learning such as long short-term memory, LSTM (Greff et al. 2017), it is worth investigating application of deep learning techniques in detecting fake location information. For example, given a set of sequenced historical geo-tags of an individual, LSTM can be used to model the sequential information and build deep learning-based classification methods.

12.5.3 Connecting Social Media with the Real World

In social media, people share their thoughts and emotions about events in the real world. Such events might be explicitly mentioned or implicitly referred to in their posts. For instance, some social media posts might explicitly include a link to a news article they would like to discuss whereas other posts might express the users' attitude on some events without citing those events. In many inference problems for social media data (e.g., sentiment analysis, opinion mining), it is crucial to determine the corresponding realistic events to fully understand and explain the trends and phenomena in social media (i.e., connecting social media with the real world). One example is that social media posts concerning implicit events where the absence of the implied events would clearly impede accurate analysis of the posts.

To model the real world, we can resort to public information resources such as news articles and public knowledge resources such as Wikipedia and Freebase (Bollacker et al. 2008). These resources cover a wide range of events across various aspects of life. They are also updated with new events in almost real time due to the recent advances in publication technologies, promoting these public information resources as a digital counterpart of the real world. Consequently, we can connect social media data with the real word via the reflected world of public information resources. The major technical challenges to accomplish this connection involve the ability to autonomously extract events from those public resources (e.g., news articles) and the capacity to link the information in social media to the appropriate detected events. Such challenges would require a deep analysis of the semantics of the information presented in both (e.g., the posts in social media and the events in public resources) to identify the events and connections with high accuracy. Fortunately, deep semantic understanding of such information is being actively investigated in artificial intelligence research, including natural language processing, computer vision, graph modeling and machine learning. For instance, many recent studies have shown that events in news articles can be effectively curated using deep learning techniques, a branch of machine learning that is capable of automatically inducing the underlying representations for data to achieve high extraction performance (Nguyen et al. 2016; Nguyen and Grishman 2018; Nguyen and Nguyen 2019). As these event extraction techniques can also recognize the time and locations at which the events occur, they can be beneficial for GIScientists in geographical research of populations on social media and the real world. In addition, deep learning might also present effective solutions for the problem of linking social media data with realistic events due to its recently demonstrated capacity for embedding and representation learning for various problems. Once converged, such advances in these fields of computer science might eventually offer an opportunity to connect the virtual world and real world by solving the aforementioned technical challenges.

Finally, the realistic events from public information resources enable novel semantic-based solutions to combat the problems of rumors or fake news in social media. An important property of the public resources discussed in this section is that they generally capture trustful information/events, as such information is verified by the media administrators for accuracy and correctness. This is one reason why news articles are usually slower than social media in presenting the information to the public. Consequently, if the social media information can be accurately linked and compared with the information/events in the trustful information sources, novel detection and tracking techniques can be proposed to prevent rumors and fact-check the information spread over social media. Artificial intelligence research can provide the fundamental technologies to tackle these problems, as demonstrated in recent research in natural language processing and deep learning (i.e., Yin and Roth 2018).

12.6 Conclusion

As a crucial platform for human dynamics and activities, social media content can be mined in multiple approaches to determine how individuals connect and share information as well as purposefully move across scales and resolutions (Croitoru et al. 2013; Miller et al. 2019). When social media activities are attached with locational information, these online human activities can generate tremendous electronic footprints on Digital Earth. Especially when merging with other digital overlays of authoritative data through multisource data fusion, such as land use, urban planning, and natural resources data, powerful interoperation and prediction that require both electronic footprints and digital overlays on Digital Earth become feasible with the optimal weights for combination (De Albuquerque et al. 2015; Lin et al. 2019). These digital overlays serve as the socioenvironmental context within which the geosocial media dynamics and events occur and evolve, calling for scientific cross-fertilization of many separate domains toward an integrated science of human dynamics.

In this chapter, we introduce the opportunities and challenges that human dynamics-centered social media bring to Digital Earth. We review the information diffusion of social media, the multifaced implications of social media, and some real-world cases. Social media will facilitate the prediction of human dynamics in a wide spectrum of aspects, including public health, emergency response, decision making and social equity promotion. Social media will also bring unintended challenges for Digital Earth, such as rumors and fake location spoofing. Considering the multifaceted implications, this chapter calls for GIScientists to raise their awareness of the complex impacts of social media and urges them to model the geographies of social media as well as filter fake locations through geographic knowledge, targeting a more robust geosocial knowledge discovery. Social media will continue to evolve, along with the development of human society. Social media has become a crucial part of human activities on Digital Earth. Any effort that ignores the importance of social media will bring the effort into question. Therefore, the study of social media provides new data sources and data collection methods about real-world activities and happenings, and social media help us in profoundly understanding ourselves as a unique species living both on the Earth and in Digital Earth.

Acknowledgements This material is partially based upon work supported by the National Science Foundation under Grant Nos. 1416509, 1535031, and 1739491. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

References

Ansari A (2012) The role of social media in Iran's green movement (2009-2012). Glob Media J Aust Ed 12:1–6

Barabási A-L (2005) The origin of bursts and heavy tails in human dynamics. Nature 435(7039):207–211

Batty M (2013) Big data, smart cities and city planning. Dialogues Hum Geogr 3(3):274–279

Bollacker K, Evans C, Paritosh P et al (2008) Freebase: a collaboratively created graph database for structuring human knowledge. In: Proceedings of the 2008 ACM SIGMOD international conference on management of data - SIGMOD '08. ACM Press, New York, NY, p 1247–1250

Chen X, Elmes G, Ye X et al (2016) Implementing a real-time twitter-based system for resource dispatch in disaster management. GeoJournal 81(6):863–873

Choudhury MD, Gamon M, Counts S et al (2013) Predicting depression via social media. In: Proceedings of the seventh international aaai conference on weblogs and social media. p 128–137

Chun Y, Kwan M-P, Griffith DA (2019) Uncertainty and context in GIScience and geography: challenges in the era of geospatial big data. Int J Geogr Inf Sci 33(6):1131–1134

Croitoru A, Crooks A, Radzikowski J et al (2013) Geosocial gauge: a system prototype for knowledge discovery from social media. Int J Geogr Inf Sci 27(12):2483–2508

de Albuquerque JP, Herfort B, Brenning A et al (2015) A geographic approach for combining social media and authoritative data towards identifying useful information for disaster management. Int J Geogr Inf Sci 29(4):667–689

Deng C, Lin W, Ye X et al (2018) Social media data as a proxy for hourly fine-scale electric power consumption estimation. Environ Plan A Econ Space 50(8):1553–1557

Finch KC, Snook KR, Duke CH et al (2016) Public health implications of social media use during natural disasters, environmental disasters, and other environmental concerns. Nat Hazards 83(1):729–760

Future of Privacy Forum and DataGuidance (2018) Comparing privacy laws: GDPR v. CCPA. https://fpf.org/wp-content/uploads/2018/11/GDPR_CCPA_Comparison-Guide.pdf. Accessed 12 May 2019

Goodchild MF (2013) The quality of big (geo)data. Dialogues Hum Geogr 3(3):280–284

Greff K, Srivastava RK, Koutnik J et al (2017) LSTM: a search space odyssey. IEEE Trans Neural Netw Learn Syst 28(10):2222–2232
Hägerstraand T (1970) What about people in regional science? Pap Region Sci 24(1):7–24
Hashimoto Y, Ohama A (2014) The role of social media in emergency response: the case of the great East Japan earthquake. NIDS J Def Secur 15:99–126
Ihde D (1990) Technology and the lifeworld: from garden to earth (No. 560). Indiana University Press, Bloomington
Jones NM, Thompson RR, Schetter CD et al (2017) Distress and rumor exposure on social media during a campus lockdown. Proc Natl Acad Sci USA 114(44):11663–11668
Kekulluoglu D, Kokciyan N, Yolum P (2017) Preserving privacy as social responsibility in online social networks. ACM Trans Internet Technol 18(4):1–22
Li M, Dong L, Shen Z et al (2017) Examining the interaction of taxi and subway ridership for sustainable urbanization. Sustainability 9(2):242
Lin J, Wu Z, Li X (2019) Measuring inter-city connectivity in an urban agglomeration based on multi-source data. Int J Geogr Inf Sci 33(5):1062–1081
Liu Q, Wang Z, Ye X (2018a) Comparing mobility patterns between residents and visitors using geo-tagged social media data. Trans GIS 22(6):1372–1389
Liu X, Macedo J, Zhou T et al (2018b) Evaluation of the utility efficiency of subway stations based on spatial information from public social media. Habitat Int 79:10–17
McLuhan M (1975) McLuhan's laws of the media. Technol Cult 16(1):74–78
Miller HJ, Dodge S, Miller J et al (2019) Towards an integrated science of movement: converging research on animal movement ecology and human mobility science. Int J Geogr Inf Sci 33(5):855–876
Nguyen TH, Grishman R (2018) Graph convolutional networks with argument-aware pooling for event detection. In: The association for the advancement of artificial intelligence (AAAI). AAAI Press, Menlo Park, California, p 5900–5907
Nguyen TM, Nguyen TH (2019) One for all: neural joint modeling of entities and events. In: The association for the advancement of artificial intelligence (AAAI), arXiv.org > cs > arXiv:1812.00195
Nguyen TH, Cho K, Grishman R (2016) Joint event extraction via recurrent neural networks. In: Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, p. 300–309
Palen L, Vieweg S, Liu SB et al (2009) Crisis in a networked world. Soc Sci Comput Rev 27(4):467–480
Rosenberg M, Confessore N, Cadwalladr C (2018) How trump consultants exploited the facebook data of millions, New York Times. https://www.nytimes.com/2018/03/17/us/politics/cambridge-analytica-trump-campaign.html?module=inline. Accessed 12 May 2019
Sharma M, Yadav K, Yadav N et al (2017) Zika virus pandemic—analysis of facebook as a social media health information platform. Am J Infect Control 45(3):301–302
Shaw S-L, Tsou M-H, Ye X (2016) Editorial: human dynamics in the mobile and big data era. Int J Geogr Inf Sci 30(9):1687–1693
Shelton T, Poorthuis A, Zook M (2015) Social media and the city: rethinking urban socio-spatial inequality using user-generated geographic information. Landsc Urban Plan 142:198–211
Shi X, Xue B, Tsou M-H et al (2018) Detecting events from the social media through exemplar-enhanced supervised learning. Int J Digit Earth 12(9):1083–1097
Siła-Nowicka K, Vandrol J, Oshan T et al (2016) Analysis of human mobility patterns from GPS trajectories and contextual information. Int J Geogr Inf Sci 30(5):881–906
Sui D, Goodchild M (2011) The convergence of GIS and social media: challenges for GIScience. Int J Geogr Inf Sci 25(11):1737–1748
Tuan Y-F (2003) On human geography. Daedalus 132(2):134–137
Villar ME, Marsh E (2018) Social media and infectious disease perceptions in a multicultural society. In: Villar ME, Marsh E (eds) Reconceptualizing new media and intercultural communication in a networked society. IGI Global, Pennsylvania, US, p 328–350

Wang Z, Ye X (2018) Space, time, and situational awareness in natural hazards: a case study of hurricane sandy with social media data. Cartogr Geogr Inf Sci 46(4):334–346

Wang Z, Ye X, Tsou M-H (2016) Spatial, temporal, and content analysis of Twitter for wildfire hazards. Nat Hazards 83(1):523–540

Yang X, Ye X, Sui DZ (2016) We know where you are. Int J Appl Geospat Res 7(2):61–75

Yao XA, Huang H, Jiang B et al (2019) Representation and analytical models for location-based big data. Int J Geogr Inf Sci 33(4):707–713

Ye X, He C (2016) The new data landscape for regional and urban analysis. GeoJournal 81(6):811–815

Ye X, Lee J (2016) Integrating geographic activity space and social network space to promote healthy lifestyles. SIGSPATIAL Spec 8(1):20–33

Ye X, Liu X (2019) Introduction: cities as social and spatial networks. In: Ye X, Liu X (eds) Cities as spatial and social networks. Springer, Cham, p 1–8

Ye X, Li S, Sharag-Eldin A et al (2017) Geography of social media in public response to policy-based topics. In: Ye X, Li S, Sharag-Eldin A et al (eds) Geospatial data science techniques and applications. CRC Press, Boca Raton, US, p 221–232

Ye X, Li S, Yang X et al (2018a) The fear of ebola: a tale of two cities in China. In: Ye X, Li S, Yang X et al (eds) Big data support of urban planning and management. Springer, Cham, p 113–132

Ye X, Sharag-Eldin A, Spitzberg B et al (2018b) Analyzing public opinions on death penalty abolishment. Chin Sociol Dialogue 3(1):53–75

Yin W, Roth D (2018) TwoWingOS: a two-wing optimization strategy for evidential claim verification. In: Proceedings of the 2018 conference on empirical methods in natural language processing. Association for Computational Linguistics, Brussels, Belgium, p 105–114

Yin J, Lampert A, Cameron M et al (2012) Using social media to enhance emergency situation awareness. IEEE Intell Syst 27(6):52–59

Zadeh AH, Zolbanin HM, Sharda R et al (2019) Social media for nowcasting flu activity: spatio-temporal big data analysis. Inf Syst Front 21(4):743–760

Zhang J, Goodchild MF (2002) Uncertainty in geographical information. CRC Press, Boca Raton, FL

Zhao B, Sui DZ (2017) True lies in geospatial big data: detecting location spoofing in social media. Ann GIS 23(1):1–14

Zhao B, Sui DZ, Li Z (2017) Visualizing the gay community in Beijing with location-based social media. Environ Plan A 49(5):977–979

Zhao B, Zhang S (2018) Rethinking spatial data quality: pokémon go as a case study of location spoofing. Prof Geogr 71(1):96–108

Zubiaga A, Aker A, Bontcheva K et al (2018) Detection and resolution of rumours in social media. ACM Comput Surv 51(2):1–36

Xinyue Ye is an Associate Professor in the College of Computing at New Jersey Institute of Technology where he directs Urban Informatics & Spatial Computing lab. He develops and implements new methods on spatiotemporal-social network analysis/modelling/simulation for different application domains such as economic development, disaster response, land use, public health, and urban crime.

Bo Zhao is an Assistant Professor in the Department of Geography at University of Washington. His research interests are on the social implications of geospatial technologies, the study of different spoofing phenomena in geography towards a post-truth reflection on the value of geospatial technologies, the geovisualization and geo-narrative of the coupled human and natural system.

Thien Huu Nguyen is an Assistant Professor in the Department of Computer and Information Science at University of Oregon. His research explores mechanisms to understand human languages for computers so that computers can perform cognitive language-related tasks for us.

Shaohua Wang is an Assistant Professor in the College of Computing at New Jersey Institute of Technology where he directs Software Engineering lab. His research spans the study of software engineering from system, empirical, and machine learning perspectives.

Part II
Digital Earth for Multi-domain Applications

Chapter 13
Digital Earth for Sustainable Development Goals

Graciela Metternicht, Norman Mueller and Richard Lucas

Abstract Sustainable development is nothing new, but it has proven notoriously difficult to implement in practice. The 2030 Agenda for Sustainable Development, with 17 goals, 169 targets and 232 associated indicators, was approved at the 2015 UN General Assembly and addresses the economic, social and environmental pillars of development, aspiring to attain by 2030 a sustainable future that balances equitable prosperity within planetary boundaries. While the goals are universal (i.e., applicable to both developing and developed countries), it is left to individual countries to establish national Sustainable Development Goal (SDG) targets according to their own priorities and level of ambition in terms of the scale and pace of transformation aspired to.

Keywords Sustainable development goals · Digital Earth · Earth Observation · Big Earth data · Indicators · Land cover classification

13.1 Fundamentals of Digital Earth for the Sustainable Development Goals

The Digital Earth (DE) exists in parallel to the physical Earth along with some translating elements between them (Sudmanns et al. 2019). Chapter 1 describes the origin, evolution and main elements of Digital Earth, and the links between Digital Earth, Big Data (Chap. 9) and big Earth data. Guo (2017) argues that, from the perspective of big data, big Earth data inherits big data's 'Vs' (volume, velocity and

G. Metternicht (✉)
School of Biological, Earth and Environmental Sciences, The University of New South Wales, High Street, Kensington, NSW 2052, Australia
e-mail: g.metternicht@unsw.edu.au

N. Mueller
Geoscience Australia/Digital Earth Australia (DEA), GPO Box 378, Canberra, ACT 2601, Australia

R. Lucas
Department of Geography and Earth Sciences, Aberystwyth University, Aberystwyth, Ceredigion, Wales SY233DB, UK

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_13

variety) and, in this context, DE can be considered to be big Earth data. Furthermore, as big Earth data research focuses on synthesis of systematic observations of the Earth, as well as data-intensive methods for studying Earth system models, based on the premise of increased knowledge discovery (Chap. 1), Digital Earth can support countries in their implementation of the 2030 Agenda for Sustainable Development.

Through analysis of recent literature and a case study, this chapter collects and presents evidence of the potential and limitations of Digital Earth for systematic generation of information and knowledge for use in measuring progress towards the Sustainable Development Goals (SDGs). We frame the analysis and discussion around priorities for implementation (ICSU, ISSU 2015), including:

(a) the design of SDG indicator metrics at national levels and how Digital Earth, through the Analysis Ready Data (ARD) concept, can contribute to that end
(b) harmonized national metrics for SDG implementation, including for baseline determination and target setting
(c) setting up monitoring platforms for tracking progress towards the SDGs
(d) knowledge needs for assessing implementation of actions and strategies towards achieving set SDG targets
(e) governance and institutional arrangements, including multi-stakeholder participation.

The remainder of the chapter is structured as follows. Section 13.2 identifies the information needs of countries for the implementation of the SDGs, including for the SDG Global Indicator Framework (GIF). Section 13.3 summarizes the findings of recent research and practice on the use of Digital Earth (including Earth Observation[1] and social sensing) in support of the SDGs. Section 13.4 presents a national case study of multi-stakeholder engagement in the operationalization of the Indicator Framework of the Sustainable Development Goals with Earth Observations. The chapter closes (Sect. 13.5) with an outlook on the prospects of Digital Earth and big Earth Data in relation to the SDGs.

13.2 Information and Knowledge Relevant to National Implementation of the SDGs

The SDGs provide a coherent, evidence-based framework for development planning and programming at a national level (Allen et al. 2017a). The goals and targets essentially set the desired destination for development through to 2030 and provide a framework for monitoring progress. This section introduces the metrics agreed for monitoring and reporting of the SDGs, and broadly identifies data and information requirements for their implementation.

[1] The Earth Observation data in this chapter refers to the definition provided by Nativi et al. (2019).

13.2.1 How the SDGs Are Monitored and Reported

The Global Indicator Framework (GIF) was established in March 2016 to monitor progress towards achieving the SDGs (UN Statistical Commission 2016). The SDG indicators have been grouped into three different tiers according to the level of data availability and methodological development. Of the 232 SDG indicators that make up the GIF, as of March 2019, 101 are classified as being Tier I. This means that the indicator is conceptually clear, has an internationally established methodology and standards, and the data are regularly compiled for at least 50% of participating countries. The remaining indicators are Tier II (94 indicators), which are conceptually clear but for which the data are not regularly produced by participating countries, or Tier III (34 indicators), for which no internationally established methodology or standards are yet available. Six indicators are determined as having several tiers (Inter-Agency and Expert Group on Sustainable Development Goals 2019). Hence, three years after the adoption of the GIF, less than half (44%) of the SDG indicators can be confidently populated.

It is worth noting that the SDG indicators are essentially **performance metrics** and, as such, are reported regularly at *national levels* through National Voluntary Reports (NVRs) (UNGA 2015, paras. 79 and 84), and annually at the *global level*. The latter is undertaken by the UN Secretary General to inform the High-Level Political Forum based on a selection of indicators from the GIF for which data are available, as mandated by the General Assembly (UNGA 2015, para. 83). For Tier I and II indicators, the availability of data at national levels may not necessarily align with the global tier classification, and countries can create their own tier classification for implementation.

13.2.2 Information Needs for Implementation of the SDGs

Recent research (Allen et al. 2018, 2019) has identified challenges for implementing the SDGs that, in turn, influence information and knowledge needs.

(a) The comprehensiveness of scope makes prioritization essential.
(b) The goals are integrated, with very complex feedback and dynamics. This is a significant change from prior narrow, linear approaches to development.
(c) The SDG targets have complex trade-offs and synergies, and conflict can emerge from the interactions between targets and goals (Lusseau and Mancini 2019; Nilsson et al. 2016; Le Blanc 2015; Allen et al. 2019).
(d) Currently, there is a weak conceptual understanding of these interlinkages, which limits the ability to respond with coherent policy and management across sectors (Allen et al. 2018; Spangenberg 2017).

Challenges related to aspects of target-setting are that the system of SDGs is not coherent, but rather a network of interlinked targets and a reflection of the political mapping of development priorities rather than a reflection of how the Earth System works (Le Blanc 2015). Furthermore, the SDGs do not reflect the cause–effect relationships that are needed to understand how the achievement of one target could impact on the other targets. Hence, national implementation of the SDGs requires more than information on performance metrics. For example, timely data in support of policy formulation and targeted interventions may be of much greater importance for countries aiming to advance the implementation of the SDGs according to their national circumstances than simply providing a metric around an agreed global indicator. Furthermore, implementation of the SDGs at national levels also requires determining a baseline for 2015, deciding on targets for 2030, as well as a system for tracking the progress towards the set targets, monitoring the performance of decisions (actions, policies and strategies) and reporting advances using the GIF.

Building an evidence-based framework for national implementation, monitoring and reporting of the SDGs requires government agencies (including National Statistics Offices) to address the what, why and how of data and information provision (Fig. 13.1).

(a) What is happening requires baseline assessment of indicators related to SDG targets, identifying priorities (e.g., what SDG targets or goals a country is lagging behind) and the identification of data and information gaps needed for such assessment, as summarized in Allen et al. (2017b).
(b) Why it is happening (e.g. drivers of and pressures leading to (un)sustainable development) relates to the need for systems analysis of interlinkages between SDG targets, understanding of cause–effect relationships, feedbacks and dynamics, and the identification of leverage points for actions and strategies to accomplish the transformational changes that the SDGs aim for.
(c) How to accomplish changes, demands that countries answering the above questions also understand how data and information are to be obtained and integrated.

Fig. 13.1 National implementation of the SDGs requires evidence-based approaches for monitoring and reporting. As implementation will largely rely upon national action, government actions, through their policy, planning, regulatory and expenditure functions—i.e. the 'plan, do, check, act' planning cycle are central to the delivery

13.3 State of the Art for the SDGs in DE

Whether 3Vs, 5Vs (including volatility and veracity, as suggested by Hammer et al. 2017) or 6Vs (including volatility, veracity and value: Fig. 13.2), big data may offer new cost-effective or efficient ways of compiling indicators, improving timeliness, and compiling linkable datasets, and also open the way for cross-cutting analyses that may help with better understanding of the causation and identification of relevant and coherent policy interventions (see Fig. 13.1).

When adopting the SDGs, the United Nations (UN) Assembly recognized the contribution that could be made by Earth Observation (EO) and geospatial information (i.e., big Earth data) in supporting and tracking progress towards the SDGs (UNGA 2015, para. 76). Analysis and interpretation of big Earth data, including Earth Observation, have much to offer the SDGs and other multi-lateral environmental agreements (Sudmanns et al. 2019). However, MacFeely (2019) makes a case for the challenges that big data face (legal, technical and ethical) concerning their use in compiling SDG indicators. National statistical offices, government agencies and UN agencies, which are the custodians of specific SDGs tasked with implementing the

Fig. 13.2 The 6Vs of big data for official statistics. Modified from Hammer et al. (2017)

GIF, face concerns about whether big data are representative and stable enough to be used consistently for compiling the SDG indicators and also their operationalization. For example, in the Big Data Project Inventory compiled by the UN Global Working Group on Big Data, 34 national statistical offices from around the world registered 109 separate big data projects and their potential contribution to the SDG implementation. Most data projects focus on goals 3, 8, 11 and, with a lesser emphasis, goals 2, 15 and 16. Though promising, most projects have not yet moved beyond the planning stage, and others are dealing with legal issues related to data protection (MacFeely 2019). Specific to the EO community are challenges for consistently and systematically turning satellite and other remote sensing data into valuable global information layers in support of effective implementation of the SDGs.

In late 2018, the Committee on Earth Observation Satellites (CEOS) compiled a report on the potential of satellite EO for the SDGs (Paganini et al. 2018), and their findings suggest that EO data has a role to play in quantifying around 40 of the 169 Targets, and around 30 of the 232 Indicators. The CEOS argues that there is an unrealized potential for EO data to contribute to the Indicator Framework, with only a third of its data being routinely exploited today. This is based on the premise that only 12 out of the 30 indicators identified are listed as Tier I.

Moreover, the report points to the importance of EO in relation to Goal 6 (Clean water and sanitation), Goal 11 (Sustainable cities), Goal 14 (Life below water), and Goal 15 (Life on land). Most of the perceived contribution of EO towards these goals has been around the provision of information in relation to the mapping of land cover, land productivity, above ground biomass, soil moisture content, and water extent or quality characteristics, as well as air quality and pollution parameters (Table 13.1). A 2016 compilation of the Group on Earth Observation (GEO) appraised the potential of EO and geospatial information for informing all SDGs, although the document was vague in terms of specific contributions to SDG targets and indicators. A subsequent joint GEO-CEOS report (CEOS-GEO EO4SDG 2017) further investigated the potential of big Earth Data (EO and geo-information) for supporting countries in the implementation of the 2030 Agenda for Sustainable Development, arguing that it could contribute to the implementation of 29 indicators (through direct measurement or indirect support) and 71 targets of 16 goals (but not all indicators of these targets). By referencing national-scale satellite datasets (e.g. Terra/Aqua MODIS, Landsat, and Sentinel), Metternicht et al. (2018) concluded that EO satellite-derived information tends to have a more indirect contribution to the SDG targets and indicators (i.e. use as proxies). Using data available from the Australian Terrestrial Ecosystem Research Network platform (TERN), the study ascertained that EO-derived information was most relevant to Goal targets 15, 14, 13, 11, 6, 3, 2 and 1, and, to a lesser extent, Goal 9 (Fig. 13.3).

The potential of EO to support the SDG indicator framework appears in the biosphere cluster (Fig. 13.4) and to a lesser the SDG indicators related to society and the economy. This concurs with the argument of Plag and Jules-Plag (2019) that very few indicators can currently be quantified based on information extracted from EO alone because of the strong focus of the SDGs on human needs and the bias toward social and economic information and the built environment. Traditional

Table 13.1 Summary of research that shows the contribution of big Earth Data towards the SDGs

SDG	Supported targets									Supported indicators				
1	1.4	1.5								1.4.2				
2	2.3	2.4	2.c							2.4.1				
3	3.3	3.4	3.9	3.d						3.9.1				
4														
5	5.a									5.a.1				
6	6.1	6.3	6.4	6.5	6.6	6.a	6.b			6.3.1	6.3.2	6.4.2	6.5.1	
7	7.2	7.3	7.a	7.b						7.1.1				
8	8.4													

(continued)

Table 13.1 (continued)

SDG	Supported targets									Supported indicators				
	9.1	9.2	9.3	9.4	9.5	9.a				9.1.1	9.4.1			
	10.6	10.7	10.a											
	11.1	11.3	11.4	11.5	11.6	11.7	11.b	11.c		11.3.1	11.6.2	11.7.1		
	12.2	12.4	12.8	12.a	12.b					12.1.1				
	13.1	13.2	13.3	13.b						13.1.1				
	14.1	14.2	14.3	14.4	14.6	14.7	14.1			14.3.1	14.4.1	14.5.1		
	15.1	15.2	15.3	15.4	15.5	15.7	15.8	15.9		15.1.1	15.2.1	15.3.1	15.4.1	15.4.2
	16.8													
	17.2	17.3	17.6	17.7	17.8	17.9	17.16	17.17	17.18	17.6.1	17.18.1			

From Paganini et al. (2018)

Product	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Forest Cover						T6.6								T14.1	T15.1, T15.3, T15.2		TERN Landscape governance as good practice of 'means of implementation'
Woody extent *															T15.1,T15.4, T15.8		
Fractional cover	T1.5/1.5.4		T3.9 / 3.d			T6.6					T11.3, T11.7		T13.1, T13.3	T14.1	T15.3, T15.1, T15.2		
Presistent Green Vegetation Fraction						T6.6								T14.1	T15.3, T15.1, T15.2		
Foliage Projective Cover						T6.6								T14.1	T15.3, T15.1, T15.2		
Seasonal Fractional Cover			T3.9 / 3.d											T14.1	T15.1,T15.3,T 15.4,T15.5		
Seasonal Ground Cover														T14.1	T15.1, T15.2, T15.3, T15.4		
Seasonal Persistent Green Cover			T3.9 / 3.d											T14.1	T15.2		
Biomass		T2.3, T2.4	T3d			T6.6								T14.1	T15.1,T15.4, T15.8		
Vegetation height and structure															T15.1,T15.4, T15.8		
Phenology		T2.4, T2a	T3d												T15.1,T15.2		
Leaf Area Index			T3d			T6.6					T11.3, T11.7		T13.1, T13.3		T15.1,T15.2, T15.3,15.5		
Gross Primary Productivity	T1.5/1.5.4	T2.3, T2.4	T3d										T13.1, T13.3		T15.3, T15.1, T15.2		
fPAR			T3d			T6.6									T15.1,T15.2, T15.3,15.5		
Active Crop Mapping*		T2.3, T2.4, T2.a											T13.1, T13.3				
Grassland curing			T3d												T15.1		
Water Prevalence	T1.5/1.5.4					T6.4, T6.6							T13.1, T13.3				
Land surface temperature	T1.5/1.5.4										T11.7			T14.1			
Thermal Anomalies	T1.5/1.5.4												T13.1, T13.3				
Burnt area & day of burn	T1.5/1.5.4												T13.1, T13.3				
NRT Burnt Area	T1.5/1.5.4												T13.1, T13.3				
Fire patchniess (sub-pixel)	T1.5/1.5.4												T13.1, T13.3				
Annual Fire Scars	T1.5/1.5.4												T13.1, T13.3				
Total Land Cover (Land Condition Index)			T3.9 / 3.d										T13.1, T13.3	T14.1, T14.2	T15.1, T15.3		
Dynamic Land Cover Map		T2.3, T2.4, T2a	T3d			T6.6			T9.1		T11.3, T11.7			T14.1, T14.2	T15.1, T15.3		
Ecosystem Disturbance Index	T6.6		T3d			T6.6									T15.1,T15.4,T 15.5,15.8,15. 9		
Surface Reflectance			T3d			T6.6								T14.1	T15.3,T15.1		
NDVI			T3d			T6.6					T11.3, T11.7		T13.1, T13.3		T15.1,T15.2, T15.3,15.5		
EVI			T3d			T6.6					T11.3, T11.7		T13.1, T13.3		T15.1,T15.2, T15.3,15.5		

Fig. 13.3 SDG targets that TERN Auscover products contribute to are listed in the table; the table cells are color-coded according to whether the contribution is more direct (green) or more indirect (yellow) (Metternicht et al. 2018)

EO techniques are designed for extracting information on environmental variables, with only a few being related to the built environment and associated infrastructure (e.g., built-up areas and roads). Hence, there are limitations on the possibility of EO *alone* producing reliable metrics for SDG indicators (see Table 13.1); however, approaches underpinned by big Earth data do have some potential, as evidenced in recent research by Kussul et al. (2019), Foody et al. (2019), Freire et al. (2018), and Corbane et al. (2017). Specifically:

- meta-optimization of EO with external data-intensive infrastructure has led to improved mapping of built-up areas in support of the global human settlement layer (Corbane et al. 2017)
- national mapping of SDG indicators 15.1.1, 15.3.1 and 2.4.1 has been achieved through synergy of in situ and multi-resolution satellite data (Kussul et al. 2019)
- big Earth Data (global census data and satellite-derived built-up area maps) has enabled enhanced mapping of population distribution along coastlines (Freire et al. 2018)
- EO and machine learning have enabled mapping of sites associated with slavery, in support of SDG target 8.7 ("*take immediate and effective measures to eradicate forced labour, end modern slavery and human trafficking and secure the prohibition and elimination of the worst forms of child labour*") (Foody et al. 2019).

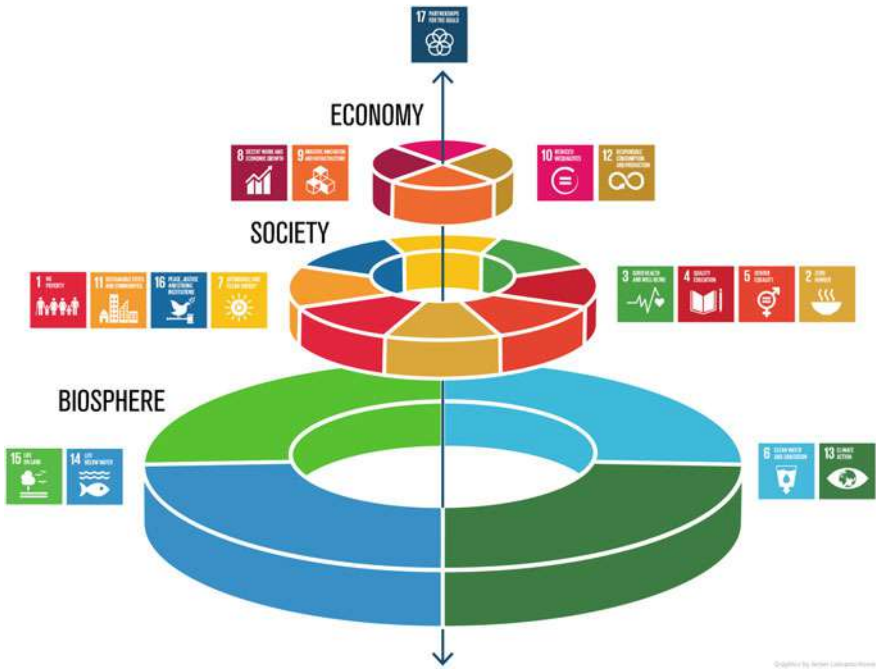


Fig. 13.4 Clustering of the SDGs that relate to the biosphere (earth life supporting system), society and economy. *Illustration* Azote for Stockholm Resilience Centre, Stockholm University

In summary, EO data does not directly deliver the SDG indicators agreed by the Inter-Agency and Expert Group (IAEG) on SDGs; rather, it provides a diversity of spatio-temporal information that can then be related to the indicator framework. For example, directly observed indicators can be specific biophysical aspects of entities (e.g., land cover status and type) that provide evidence for monitoring advances towards SDG targets. As an example, changes in land-cover states can be an indication of land improvement or land degradation in SDG target 15.3. Indirect cues derived from EO data can provide evidence for SDG domains related to human health, cities and infrastructure, ecosystem health and so on (Paganini et al. 2018; Sudmanns et al. 2019). Few studies, however, refer to specific SDG indicator metrics; many papers and reports highlight the potential of Earth Observation for targets and goals but fall short of being specific regarding the operationalization of Digital Earth for the SDG target or indicator.

For the full information potential of big EO data in support of the SDGs to be realized, approaches are needed that broaden the use of EO beyond specialized scientific communities and that support decision makers with the knowledge required by systematically analyzing all available observations by converting them into meaningful geophysical variables. Data Cubes (see Chap. 21) apply the concept of satellite ARD and are facilitating access to large spatio-temporal data (Giuliani et al. 2017). This enables the coupling of EO with other big data such as demographic, economic, climatic, or administrative data, which are needed to make indicators and analysis more relevant and targeted to the SDGs. Furthermore, some of the proposed SDG targets relate to the so-called 'means of implementation', namely technology transfer and capacity building (i.e. SDG17; SDG targets 13.1, 1.3.3 and 16.8). In this regard, Digital Earth and EO infrastructure, as currently offered by Australia's TERN Landscape initiative (TERN 2017) and other major international and national systems for big Earth data (e.g. Google Earth Engine, Amazon Web Services, Earth Server, Earth Observation Data Centre, Copernicus Data and Exploitation Platform-Deutschland, United States Geological Survey Earth Explorer, Swiss Data Cube, Digital Earth Australia, Chinese Academy of Sciences Earth, and GEO*Essential* of the Group on Earth Observations), could serve as 'methodological frameworks' and examples of good practice for cross-institutional governance models, thus indirectly contributing to progress towards these targets.

The case study presented hereafter is an example of how EO can be a promising complement to traditional national statistics. Digital Earth Australia (DEA) aligns with the current trends in EO of having open data policies and using cloud computing and data cubes for improving big Earth data integration and analysis, thereby strengthening environmental data and indicators (Dhu et al. 2017). In particular, this case shows how the analysis capabilities of DEA (see Chap. 21 for infrastructure) can be used to draw together and effectively link data from multiple domains in support of the implementation of the 2030 Agenda for Sustainable Development in Australia.

13.4 Case Study of Australia: Operationalizing the Indicator Framework of the SDGs Through DE and a Participatory Process

In July 2018, Australia produced its first Voluntary National Review (VNR) of the SDGs (Australian Government 2018). Australia's consideration of the SDG Indicators has been a whole-government exercise. The Australian Bureau of Statistics (ABS) undertook a data-mapping exercise for the SDGs, in conjunction with lead agencies, exploring both ABS and other government-held data sources to identify those germane to supporting monitoring and reporting on the SDGs. A Reporting Platform[2] was created to: (a) house identified Australian government datasets relevant to the development of the country's SDG indicator framework; (b) assist in identifying new datasets; and (c) refine the SDG indicators, particularly as the move from a Tier III to a Tier I or II occurs and where additional datasets may be needed. An inter-agency governance agreement assigned the responsibility for following up and completing additional data sets to individual agencies (particularly those that hold datasets relevant to the SDG indicator framework).

[2] https://www.sdgdata.gov.au/.

For the first VNR, a total of 118 indicators were reported online using data drawn from a national indicator dataset. For 57 indicators, potential data sources were identified. However, further analysis is needed to ensure the data are suitable for reporting and are comparable to the globally agreed methodology for each UN SDG indicator. 12 indicators were not reported either because the indicator was not relevant to Australia or because no suitable Australian government data source exists for the indicator. Another 57 were not considered because, at the time of reporting, a globally agreed methodology for these UN SDG indicators is lacking (i.e., Tier III). Therefore, Australia did not investigate potential data sources. In summary, the first Australian VNR took a narrative approach, addressing each of the SDGs, though no baseline was created. Targets were not specified and Australia had complete and relevant datasets for only half of the SDG indicators. The Australian government has acknowledged that EO technology can help progress towards the completion of datasets and, in tandem, inform decision-makers about performance against SDG targets and indicators (Australia Government 2018).

In this regard, EO-derived information could help in setting baselines against which SDG targets could be set and, in turn, measure progress against agreed goals—aspects that the first VNR did not tackle. Germane to this point is the DEA initiative led by Geoscience Australia, which has enabled the compilation, analysis and interpretation of decades of satellite sensor (largely Landsat) data into information and insights about Australia's terrestrial and marine ecosystems using ARD standards (Dhu et al. 2017; Lewis et al. 2016). Building on the DEA infrastructure (see Chap. 21), Geoscience Australia is leading an inter-institutional initiative to produce reliable, standardized, continental-scale maps of land cover and land-cover dynamics across Australia at 25 m spatial resolution using multi-scale time series of Landsat and Copernicus Sentinel datasets. This approach builds on the Earth Observation Data for Ecosystem Monitoring (EODESM; Lucas and Mitchell 2017), which is fully described in Lucas et al. (2019a) and which provides multi-scale and temporal land-cover and evidence-based change maps by integrating environmental variables retrieved from EO data and utilizing the framework of the Food and Agriculture Organisation (FAO) Land Cover Classification System (LCCS; Version 2, Di Gregorio 2016). The approach is based on the requirement for information about land cover and its change over time, as both are essential input metrics to several SDG targets (Fig. 13.3) and indicators (e.g. 6.6, 11.3.1, 15.2.1, 15.3.1). This information is also useful to other national and international reporting requirements on the state of the environment (e.g. United Nations Convention to Combat Desertification, Aichi Targets, and the Paris Agreement).

13.4.1 DEA to Map Land Cover and Dynamics Over Time

The DEA land cover product has been optimized for high-performance computing within the Open Data Cube (ODC) framework and is generating continental maps of land-cover datasets from environmental variables (thematic and continuous), with a focus on those that are generated at a national level within DEA's ODC environment (Lucas et al. 2019a) and for multiple points in time. These include the vegetation cover fraction of the Joint Remote Sensing Research Program (Gill et al. 2017), Water Observations from Space (WOfS) (Mueller et al. 2016), surface reflectance Median Absolute Deviation (MAD) (Roberts et al. 2018), and national mangrove distribution (Lymburner et al. 2019) (Fig. 13.5). Additional layers generated through DEA are also used (e.g., the InterTidal Elevation Model (ITEM) of Sagar et al. (2017). The mapping is undertaken at 25 m resolution and the initial focus has been on generating land-cover classifications according to the LCCS Level 3 taxonomy, which differentiates 8 classes relating to aquatic and terrestrial (semi) natural vegetation, cultivated and managed terrestrial and aquatic vegetation, artificial and natural (bare) surfaces, and natural and artificial water bodies (Fig. 13.5 and Table 13.2). More detailed classifications are being generated at what is termed Level 4 (e.g., vegetation canopy cover and height, and water hydroperiod), which are further described using

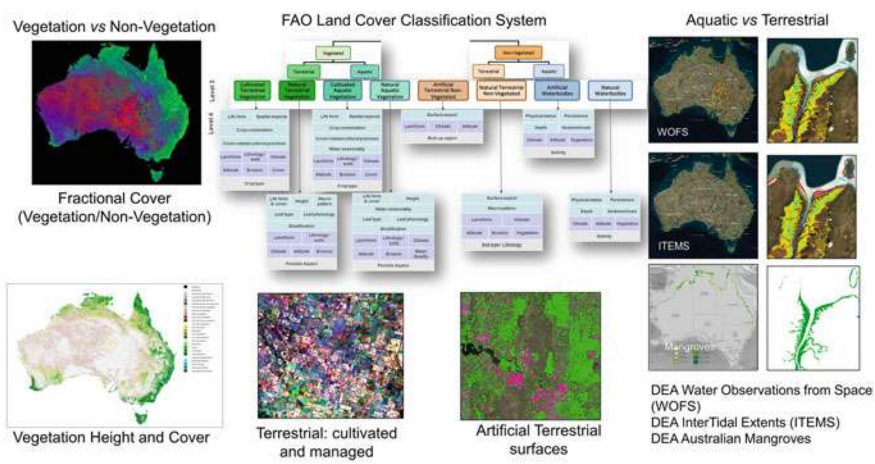


Fig. 13.5 Examples of data inputs for the application of the FAO LCCS level 3 within Digital Earth Australia used to produce standardized land cover maps at 25 m resolution

Table 13.2 Level 3 FAO land-cover classification (FAO LCCS) classes

Class name	Acronym
Cultivated terrestrial vegetation	CTV
Natural terrestrial vegetation	NTV
Cultivated aquatic vegetation	CAV
Natural aquatic vegetation	NAV
Artificial terrestrial non-vegetated	AS
Natural terrestrial non-vegetated	BS
Artificial waterbodies	AW
Natural waterbodies	NW

environmental variables that are external to the LCCS taxonomy (e.g., soil moisture and crop type), examples of which are given in Fig. 13.6.

The availability of multi-temporal land-cover layers enables change matrices (e.g. $T_1 - T_{baseline}$) to be generated between land covers obtained for any two time-separated periods. When only the LCCS Level 3 is considered, the temporal comparison between two land-cover maps results in 64 different change categories (Fig. 13.7a).

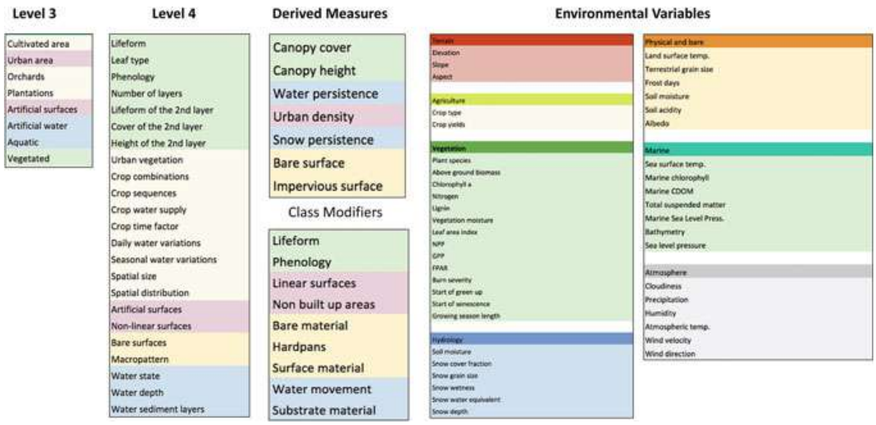


Fig. 13.6 Examples of environmental variables, class modifiers and derived measures required to implement the LCCS at level 3 and level 4 in Australia

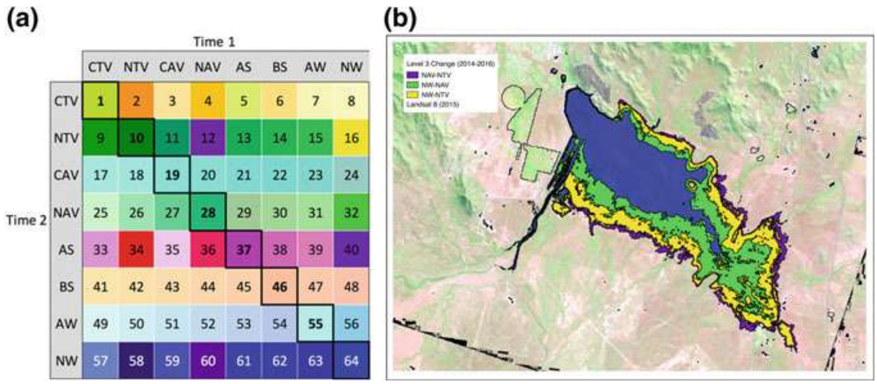


Fig. 13.7 **a** The 64 change categories generated through comparison of 2 LCCS Level 3 classifications (each with 8 classes) in the vicinity of Lake Ross (area of Townsville, Queensland) based on multi-temporal classification of Landsat images using LCCS level 3. The key changes are NAV-NTV: denoting changes from Natural Aquatic Vegetation (2014) to Natural Terrestrial Vegetation (2016); NW-NAV: Natural Waterbodies to Natural Aquatic Vegetation; and NW-NTV: Natural Waterbodies to Natural Terrestrial Vegetation. **b** The corresponding change map indicating a progressive loss of open water area, the retreat of aquatic (wet) vegetation and a transition to drier vegetation on the outer margins of the lake basin (Lucas et al. 2019a)

Diagonal cells represent areas where the land cover (e.g. natural/semi-natural terrestrial vegetation, natural water, artificial water, etc.) remains stable between the two time periods and unique codes can be assigned for the *From* → *To* changes in land cover. Figure 13.7b provides an example of a land-cover change matrix and map that result from applying FAO LCCS level 3 on an inland water ecosystem in the State of Queensland between two time periods.

One aspiration of DEA's land cover product is to better inform management and interventions in order to advance assessment and monitoring of progress towards the SDGs at national levels. In this regard, research is being undertaken to concurrently develop a change alert system (historically and when new data and data products become available) that can associate changes in states (i.e., environmental variables) with the causative mechanisms (i.e., human activities and climatic variability) and the impacts that such changes produce (e.g. defoliation, land clearing, and increases in built-up area). Such changes are based on evidence, and exploit a newly developed change taxonomy (Lucas et al. 2019b). Geoscience Australia is extending the idea to integrate, within DEA's land cover product, EODESM with the Drivers-Pressure-State-Impact-Response (DPSIR) framework (Lucas et al. 2019b; Metternicht et al. 2019). In doing so, links are—between economic and climate drivers and pressures of change and detailed information on states, state changes and environmental impacts (based on the change taxonomy). The drivers-pressure-state links can subsequently inform impacts on management and policy (from local to international l-levels). The ultimate ambition is to generate options for context-based policy and management responses related to the SDGs (Fig. 13.8). Through this approach, responsible

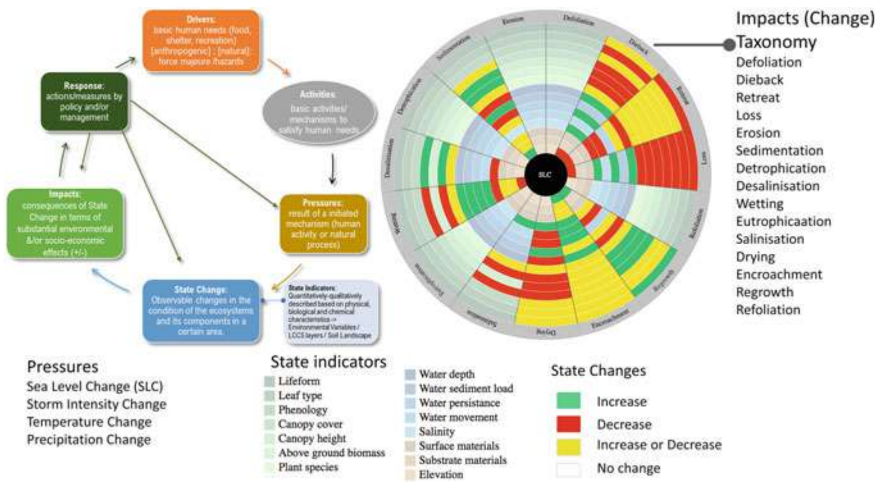


Fig. 13.8 Conceptual framework that links the DPSIR framework with the LCCS-derived land-cover maps within the DEA environment. Pressures (center of the wheel) are identified and state indicators derived from the LCCS comparison between T_0–T_1 provide an estimation of state change. Cumulative information on state change builds evidence on impacts (outer part of the wheel)

authorities can make informed and timely decisions on interventions (e.g. management decisions, new regulations).

As an illustration of the application of the integrated EODESM-DPSIR framework, Fig. 13.9a shows the impact of rising sea levels (between 1991 and 2011) on water and vegetation variables in Kakadu National Park, located in Australia's Northern Territory. An increase in water depth, salinity and hydroperiod and a corresponding rise in vegetation biomass, height and cover, along with an associated transition from shrubs to trees (i.e., lifeform state change) was observed during this period. Such changes might lead to an increase or a decrease in mangrove species. In 2015, a substantive drop in sea level in the Gulf of Carpentaria (Duke et al. 2017) was also noted in the Northern Territory (Lucas et al. 2018), which led to changes in water conditions and a substantive dieback of mangroves. A loss of canopy cover (%) and above-ground biomass (Mg ha^{-1}) were the EO-derived state-change indicators of short-term change; they were mapped through multi-temporal comparison (2014–2016) of vegetation indices (primarily a Normalized Difference Vegetation Index (NDVI) and a Plant Senescence Reflectance Index (PSRI)) derived from Rapid-Eye satellite imagery. Dieback-affected mangroves were not removed and their height (m) did not change (at least in the short term). A reduction in moisture content (%) of woody vegetation was the proxy applied to differentiate dieback from defoliation (Fig. 13.9b). Information on this proxy indicator can be discerned from, for example, time series of Advanced Land Observing Satellite (ALOS) Phased Arrayed L-band Synthetic Aperture Radar (SAR) data. Figure 13.9 shows further aerial images of sea-water encroachment along creeks and the associated colonization by mangroves (9e), as well as mangrove dieback along the eastern and western shores of the West Alligator River (9f).

The combination of the EODESM and DPSIR frameworks enables mapping of where and how much change has occurred (extent and magnitude), the root causes (sea-level change), and impacts (e.g., regrowth and dieback). Furthermore, likely impacts on policy (e.g., the United Nations Framework Convention on Climate Change or the Convention on Biological Diversity) and land management (e.g., associated with Kakadu National Park) can be indicated and future interventions suggested. In the case of SDG 6.6 ("*By 2020, protect and restore water-related ecosystems, including mountains, forests, wetlands, rivers, aquifers and lakes*"), main policy actions to advance this target should address drivers of climate change (Metternicht et al. 2018; Asbridge et al. 2018), including also environmental monitoring through Digital Earth platforms (Lymburner et al. 2019).

Ongoing research is focusing on the use of DEA's land cover product to derive Australia-wide indicators for SDGs 6.6.1 (change in the extent of water-related ecosystems over time), 11.3.1 (ratio of land consumption rate to population growth rate), 15.1.1 (forest area as a proportion of total land area) and 15.3.1 (proportion of land that is degraded compared to total land area). For example, the 2018 Australia VNR mentions that the country is 'exploring data sources' for the implementation of Indicator 15.3.1.

The following are examples of how multi-temporal land cover maps produced within DEA using ARD satellite imagery (Landsat or Sentinel) and the combined

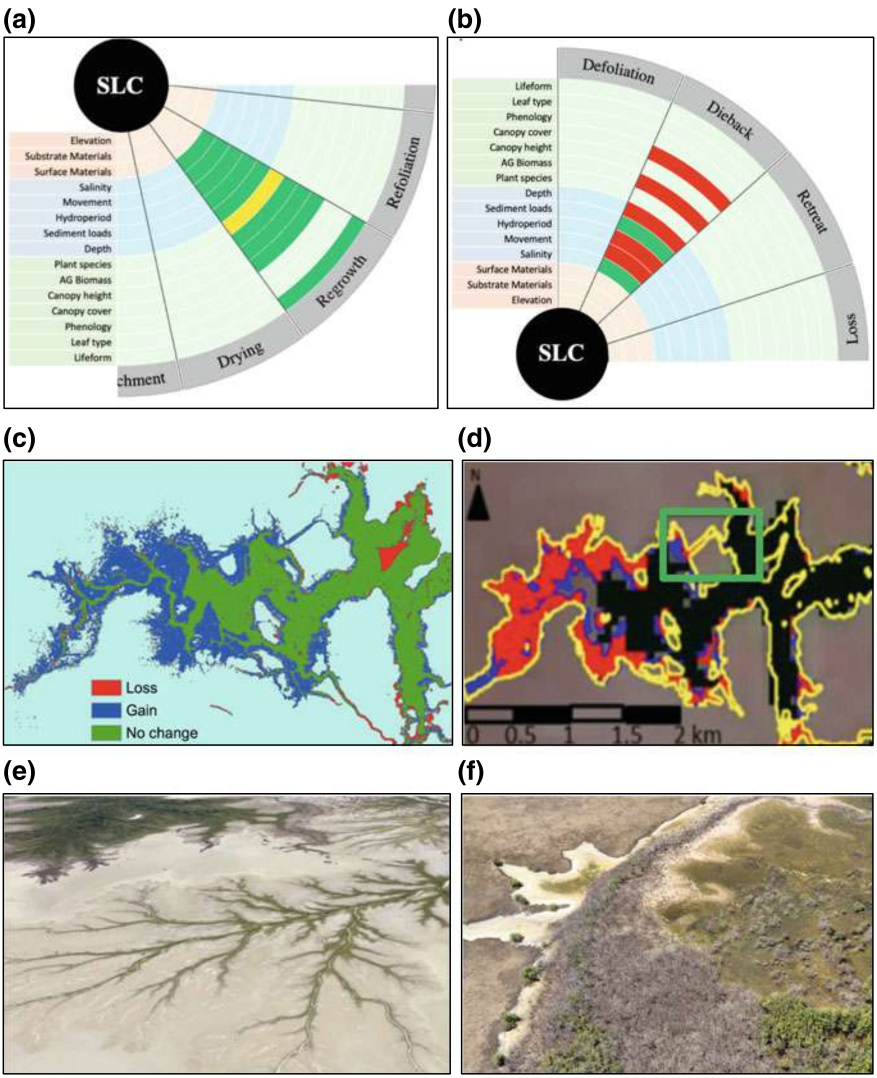


Fig. 13.9 Example of the application of the combined EODESM-DPSIR framework within DEA for Kakadu National Park, NT, Australia, where the impacts of sea-level change (SLC; center) result in **a** regrowth and colonization when rises occur and **b** dieback when drops in sea level follow. These impacts are illustrated by **c** high-resolution maps of change from time-series comparison of aerial photography from 1991 and LiDAR from 2011 (Asbridge et al. 2016), and **d** comparison of RapidEye data from 2014 and 2016. Aerial images of mangrove change taken in September 2016 show **e** landward colonization along small creeks and **f** dieback (see green box in **d**)

EODESMDPSIR framework could be used to derive metrics needed for baseline setting, target setting and/or monitoring and reporting of SDG Target 15.3, which aims '*to combat desertification, restore degraded land and soil, including land affected by desertification, drought and floods, and strive to achieve a land degradation-neutral world by 2030*'.

13.4.2 DEA in Support of SDG Indicator 15.3.1

In the SDG Global Indicator Framework, indicator 15.3.1 "Proportion of land that is degraded over total land area" is based on the analysis of available data for three sub-indicators: land cover, land productivity and carbon stocks; this indicator takes a binary form (degraded/not degraded). Computing SDG Indicator 15.3.1 for the baseline (i.e., $T_{baseline}$) and subsequent monitoring years (T_1–T_n) requires adding up all those areas where any changes in the sub-indicators (i.e. land cover, land productivity and soil organic carbon) are considered negative (or stable if the baseline or previous monitoring year labeled the area 'degraded') by national authorities. In turn this involves:

i. assessing the land cover and changes in land cover (i.e., trends)
ii. analyzing the status of and trends in land productivity based on net primary production
iii. determining carbon stock values and changes, with an initial assessment of soil organic carbon as the proxy (Sims et al. 2017).

As a proxy for measuring progress towards SDG Target 15.3, indicator 15.3.1 presupposes that changes in land cover may point to land degradation if such change implies a loss of ecosystem services considered desirable in a local or national context. Hence, land cover information at the national level derived from a classification system such as the FAO LCCS can be used to assess and quantify land cover and trends in land-cover change (Step *i* from above) by disaggregating the landscape into 'degraded/negative/declining', 'stable/unchanging' or 'improving/positive'.

Based on the example presented in Fig. 13.10, the change matrix (containing 64 possible types of land change), obtained by comparing two satellite images from two different periods classified using LCCS Level 3, can be translated into descriptors relevant to SDG indicator 15.3.1. Changes indicative of land degradation can be decided by individual countries, according to their national circumstances. In Fig. 13.10, changes highlighted in orange (e.g., agricultural and urban expansion, wetland drainage and vegetation loss) are considered examples of land degradation. Diagonal cells in blue denote areas of no change (i.e., the land cover remained stable between periods 1 and 2).[3] Cells in green denote changes that the country would

[3]It is worth noting that land degradation can still occur within classes considered stable at LCCS Level 3. For example, a landscape may remain classified as terrestrial semi-natural vegetation at both T_1–T_2 even though a loss of canopy cover may have occurred. This is described at Level 4 of the LCCS (as illustrated in Fig. 13.9).

(a)

Time 2 (*To*)

Time 1 (*From*)	CTV	NTV	CAV	NAV	AS	BS	AW	NW
CTV	Stable	Vegetation establishment	Change in agriculture	Vegetation establishment	Urban expansion	Vegetation loss	Inundation	Inundation
NTV	Agricultural expansion	Stable	Agricultural expansion	Inundation	Urban expansion	Vegetation loss	Inundation	Inundation
CAV	Agriculture drainage	Reforestation/ Abandonment	Stable	Reforestation/ Abandonment	Urban expansion	Vegetation loss	Inundation	Inundation
NAV	Wetland drainage	Wetland loss	Agricultural expansion	Stable	Urban expansion	Vegetation loss	Inundation	Inundation
AS	Withdrawal of settlements	Withdrawal of settlements	Withdrawal of settlements	Withdrawal of settlements	Stable	Withdrawal of settlements	Inundation	Inundation
BS	Agriculture expansion	Vegetation establishment	Agriculture expansion	Vegetation establishment	Urban expansion	Stable	Inundation	Inundation
AW	Agricultural expansion	Vegetation encroachment	Agricultural expansion	Vegetation encroachment	Urban expansion	Drought conditions	Stable	Removal of infrastructure
NW	Wetland drainage	Vegetation encroachment	Agricultural expansion	Vegetation encroachment	Urban expansion	Drought conditions	Infrastructure development	Stable

(b)

Count	ChangeType	Label	Description	Percentage
7281826	stable	NTV	stable	45.511412
1922451	degradation	NTV -> AS	urban expansion	12.015319
1661707	stable	NW	stable	10.385669
632635	degradation	NTV -> BS	vegetation loss	3.953969
551990	degradation	NAV -> NTV	wetland loss	3.449938
517389	stable	CAV	stable	3.233681
503453	improvement	BS -> NTV	vegetation establishment	3.146581
492064	stable	BS	stable	3.0754
374417	degradation	BS -> AS	urban expansion	2.340106
354845	stable	AS	stable	2.217781
315979	improvement	CAV -> NAV	vegetation establishment	1.974869

CTV Cultivated Terrestrial Vegetation
NTV Natural Terrestrial Vegetation
CAV Cultivated Aquatic Vegetation
NAV Natural Aquatic Vegetation
AS Artificial Surfaces
NS Natural Surfaces
AW Artificial Waterbodies
NW Natural Waterbodies

Fig. 13.10 Example of deriving the sub-indicator 'trend in land cover' through a change matrix that compares land-cover changes from time 1 to time 2. The land-cover layers are produced using the FAO LCCS level 3 and EO ARD available within the DEA. Expert knowledge input is needed to decide whether a change *From To* expresses an improvement (green cells), stability (blue cells), or degradation (orange cells)

consider to correspond to a decrease in degraded areas (i.e., an improvement) as a consequence of, for instance, sustainable land-management interventions that were made during the time period T_1–T_2. Figure 13.10b shows the output of this EO-based mapping process, summarizing the number of hectares of land that remained stable, were improved or have been degraded further between T_1 and T_2. This output can then be overlain and integrated with national information on land productivity status and trends, as well as soil organic carbon stocks, as suggested by the GIF metadata and good practice guidance for Indicator 15.3.1 (Sims et al. 2017).

Although it is still at the proof-of-concept stage, these applications show the potential of Digital Earth to assist countries in meeting several of the SDGs (particularly 6.6, 13, 14, and 15) where land cover and its change dynamics are relevant to reporting on the approved indicator (metric), tracking progress towards their attainment by 2030, helping to set targets according to national circumstances, and importantly, setting baselines. The baseline year for the SDG indicators is 2015 and for those related to land, its value (t_0) should be derived from time-series data for the period

2000–2015. The retrospective capacity of data provision by EO provides a unique comparative advantage to the achievement of this ambition.

13.4.3 Digital Earth in Support of SDG 17: Strengthen Means of Implementation

DEA is an example of big Earth data contributing to SDG 17 in aspects such as multi-stakeholder partnership, and production of data and systems for monitoring and accountability, and is also enhancing capacity-building support to developing and least-developed countries. The capabilities of the ODC to provide EO ARD and for scaling out across the world are significant contributions to Goal 17 in terms of strengthening means of implementation through technology transfer, capacity building and data, and monitoring and accountability.

The technology that lies beneath DEA, which was pioneered by Geoscience Australia, The Commonwealth Scientific and Industrial Research Organisation, and Australia's National Computational Infrastructure, underpins ODC initiatives being rolled out in developed (e.g. Switzerland) as well as developing countries (e.g. Vietnam) and regions (e.g. Digital Earth Africa: DEAfrica). DEAfrica is an example of Australia fulfilling Goal 17's aim of strengthening the means of implementation, as it builds technical and policy expertise as well as data analysis capability in-country with technical and operational guidance from DEA. A public–private investment partnership will provide continuing investment for DEAfrica, and it is envisaged that analysis, products and tools produced by DEAfrica will be accessible across the continent to inform decisions about land and water.

13.4.4 The Way Forward: Partnerships to Strengthen DEA in Support of the SDGs

The Australian Bureau of Statistics and other lead agencies (e.g. Department of Environment and Energy) that have contributed to the development of the Australian Reporting Platform (Fig. 13.11) recognize the importance of partnerships and collaboration with data providers for collecting datasets relevant to the SDG indicator framework. Big Earth data is needed to track the progress of Australia's performance on the goals and set targets, in addition to reporting to the United Nations High-Level Political Forum on the SDG Indicators Framework. Multi-source, multi-temporal data covering the socio-economic and environmental pillars of sustainable development can also assist in identifying interlinkages, overlaps and interactions between the SDGs, a key issue in the development of coherent policies and interventions, as discussed in Sect. 13.1.

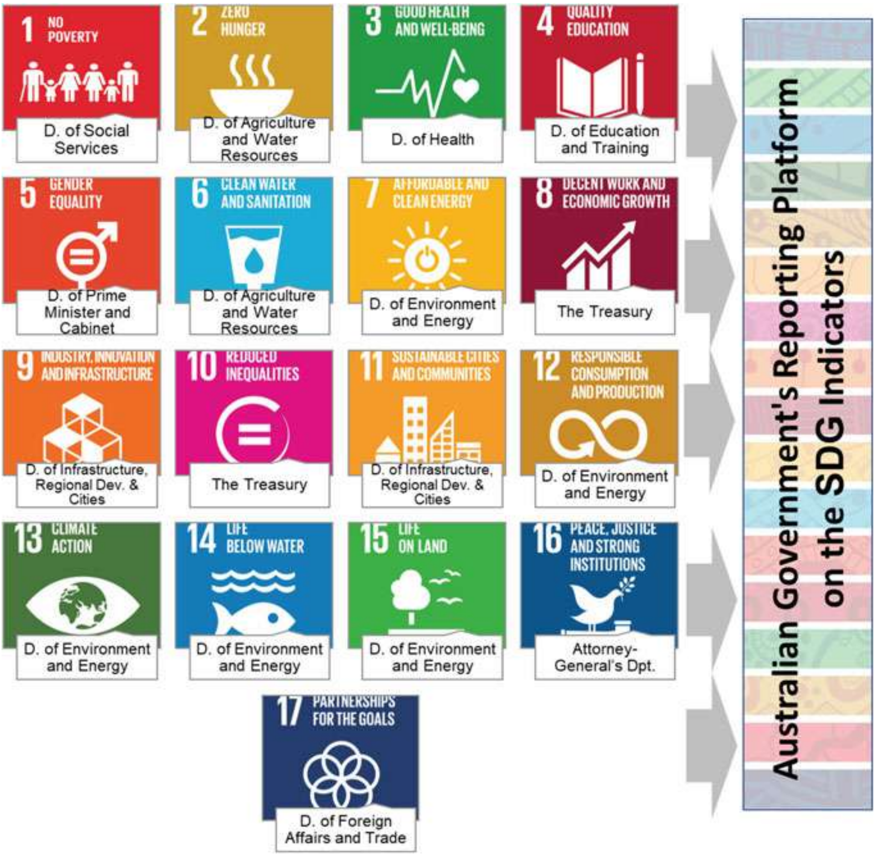


Fig. 13.11 The Australian Government's Reporting Platform for the SDGs adopts a participatory, whole-government approach

As progress is made on identifying datasets and on refining the SDG Indicators, particularly as they move from Tier III to Tier I or II, additional datasets will be uploaded to the platform, offering new data for indicator metrics and enabling the development of time-series of datasets. The government plans that the platform can assist in streamlining reporting for other nationally and internationally agreed goals (e.g. Aichi Targets, Sendai Framework, and implementation of the System of Environmental Economic Accounts (SEEA) framework). In keeping with the intention of the SDG indicator framework, the official GIF may be complemented by SDG indicators that are relevant at the regional and national levels (Australian Government 2018).

13.5 Big Earth Data for the SDG: Prospects

Measuring progress for the SDG targets through the Global Indicator Framework requires metrics that rely on biophysical, social, and economic data and information. This chapter has reviewed the current role of Digital Earth (EO as a sub-set of big Earth data) in the SDGs. It can be seen that progress has been made on identifying EO data and information for the SDG GIF (Sect. 13.3), and that participatory, cross-institutional approaches developed under a "Digital Earth" umbrella can deliver operational, standardized information that contributes to baseline and target setting, and to tracking progress towards the SDGs (4). Opportunities, and associated challenges, exist in relation to the realization of the full potential of DE for the SDG. This final section identifies and discusses these in terms of three main aspects: research and development (R&D) and technology; governance, institutional and normative aspects; and the science-policy interface.

13.5.1 R&D and Technology

Social sensing and other big data integrated within DE have the potential to meet current information and knowledge gaps for SDG indicators focused on socio-economic information (e.g. zero hunger, good health and well-being, and gender equality). Plag and Jules-Plag (2019) and Dong et al. (2019) conclude that new geospatial information for sustainability (e.g. on the built environment, land use and management), could be derived from the integration of traditional EO approaches to data gathering with citizen science, crowd-sourcing, social sensing, big data analytics and the Internet of Things. Hence, further research is needed to better establish how countries can profit from these new technologies for data gathering and analysis, embedded in a DE framework, and advance the development of indicators complementary to the core of the SDG GIF. This can support country-based interpretation and better, more coherent, narratives of national progress towards the 2030 Agenda for Sustainable Development (Metternicht et al. 2019).

Information on the *use and management* of land rather than land *cover* is needed for many SDGs (see Sect. 13.3 and Wunder et al. 2018); hence, it is relevant and pertinent to develop 'Essential Land Variables' or 'Essential Land Use Variables' to better support the information needs of the SDG targets and indicators. Digital Earth data, technology and analytics can underpin primary observations of the changes in state of land-related variables (Dong et al. 2019), with the potential to be linked to state-change indicators or to the pressures driving changes in state (see Sect. 13.4 and Lucas et al. 2019b), thus contributing to tracking progress on SDG implementation. Recent research (Plag and Jules-Plag 2019; Masó et al. 2019) has put forward ways of improving the current SDG indicator framework through considering Essential Variables. The Group on Earth Observation (GEO) and major international networks such as the Biodiversity Observation Network (GEO-BON) and the Global Ocean

Observing System (GOOS) have developed essential variables on climate (ECVs), oceans (EOVs), the water cycle (EWVs), and biodiversity (EBVs). However, standardized essential variables related to land (ELV) (or land use: ELUVs) are lacking. Global programs (e.g., Future Earth's GLP[4]) and EU-funded initiatives (e.g., the GEOEssential, ERA-PLANET[5] and ConnectinGEO projects) have started discussions on the design and development of essential land variables; the research of Reyes et al. (2017), Masó et al. (2019), Lehmann et al. (2019), Nativi et al. (2019), and Plag and Jules-Plag (2019) provide the conceptual principles and the information needs that these variables should fulfil in order to address current SDG policy and the knowledge needs of indicators. A constellation of Essential Variables on land cover/use, agriculture, biodiversity, water, and climate could better support implementation of the SDGs and the associated GIF, and also underpin systematic generation of sustainability-related knowledge from big Earth data. This would benefit Agenda 2030's global-change policy, as well as other major international agreements and conventions (e.g. the Sendai Framework for Disaster Risk Reduction, and the Paris Agreement on Climate Change).

13.5.2 Normativity, Governance and Institutional Arrangements

Google Earth Engine and Amazon Cloud-based Web Services are among cutting edge initiatives providing efficient solutions that lower the barriers to ARD products. These allow users to concentrate on data analysis and interpretation for better use of the growing volume of EO data (Giuliani et al. 2017), and expand the ecosystem of 'next users' beyond specialized scientific communities. While this is a key requirement for unlocking the informational power of big EO data and expanding the number of potential EO data users, it presents normative and governance challenges concerning big data veracity (Dong et al. 2019). Lowering access barriers for data analytics by users beyond the scientific community could potentially deliver low-quality information products. In this regard, the DE community needs to expand and build upon existing norms, standards and guidelines that have been advanced in the context of EO data storage and processing (see Sudmanns et al. 2019) to include data validation and quality assurance for information products. For example, Hernandez (2017) postulates that Digital Earth will need to consider how to store the proper metadata so that any user can easily understand how accurate data are, and how the quality of the data has been evaluated or validated. More to the point, he argues for adequate e-infrastructure and standards.

[4]An 'Essential Land Use Variables world café' session was held at the 4th Open Science Meeting of the GLP, Bern, Switzerland, April 2019. https://www.conftool.com/osm2019/index.php?page=browseSessions&cols=3&form_session=112&mode=table.

[5]ERA PLANET: The European network for observing our changing planet.

Normative challenges also remain regarding how best to determine the quality and veracity of big data from a statistical perspective (e.g., ethical questions regarding ownership of data and products). What is legally, ethically and culturally acceptable for accessing and using big data? What should the governance of digital repositories, particularly those hosting globalized or multi-national big data sets, look like? MacFeely (2019) rightly reflects that "*open cloud, centralised statistical production rather than replicating many times in countries is tempting, though it faces challenges of data and information sovereignty, as it places data owners and the data themselves beyond the reach of national level systems*".

Institutional adaptation for transformative data and information acquisition is needed as well. National Statistical Offices (NSOs) are tasked with assembling relevant data for national voluntary reports on the SDGs. The big Earth data community needs to understand how best to engage with this community to develop metrics derived from EO data that can be used for reporting. Soulard and Grenier (2018) summarize the challenges of using EO data for official statistics. Among the most salient are that datasets created from EO were not designed for use as official statistics. For integration of the EO datasets, and to better exploit the potential of big Earth data, Soulard and Grenier argue that NSOs need to develop methodologies to properly interpret existing datasets to provide estimates required by official statistics; evaluate the pertinence of global datasets that are often designed without regional considerations; keep up with the ever-increasing number of EO-generated datasets; adjust the national or regional data where local data of better quality highlight important shortcomings in the national or regional dataset; evaluate the complementarity of using EO data where other data often does not exist; and influence EO producers to integrate official statistical objectives into the EO processing workflow from the beginning. It is a two-way communication process.

13.5.3 Science-Policy Interface

Operationalization of big Earth Data proof-of-concepts is relevant to the scientific support for sustainable development policy strategies that are coordinated and coherent across goals. Reflecting on the status of operationalization of big data for SDGs from the perspective of NSOs, MacFeely (2019) argues that "*Advances, such as, the Internet of Things and biometrics will all surely present opportunities to compile new and useful statistics. The implications of this 'big (data) bang' for statistics in general, and the SDGs in particular, is not immediately clear, but one can envisage a whole host of new ways to measure and understand the human condition and the progress of development*". The UN Economic Commission for Europe (2016) reflecting on their experiences, noted 'High initial expectations about the opportunities of Big Data had to face the complexity of reality. The fact that data are produced in large amounts does not mean they are immediately and easily available for producing statistics'. Simply put, the interface between science and policy needs enhancement for context-based interpretation and communication as discussed below.

The implementation of 'transformational' policies and strategies for achieving the goals of the 2030 Agenda for Sustainable Development requires tracking the progress of set targets to ensure that responses to interventions (e.g., land restoration or sustainable cities) are as expected. In this regard, a major challenge of Digital Earth is the linking of scientific results concerning knowledge derived from EO to the policy decision space. On the one hand, multi-stake, whole-government, participatory processes, as implemented by the Government of Australia in setting its National Reporting Platform (see 4.1 and 4.4), contribute to bridging the gap between science and policy. On the other hand, DE frameworks more focused on the 'knowledge' element of the Data-Information-Knowledge-Wisdom (DIKW) paradigm are needed. SDG indicators should provide policy makers with the knowledge necessary for wise decisions, drawn from information gathered from observed data, whether through EO, social sensing, or other means. (Nativi et al. 2019). Most DE initiatives currently focus on 'Data' (i.e., ARD) as shown in the review by Sudmanns et al. (2019) of popular systems and portals for accessing or processing EO. This review makes clear that many portals facilitate data access—although in the end users struggle to produce information and 'frame' it according to context. This is an essential aspect of the policy and political decision-making processes related to the implementation of the SDGs, given that countries are to take into account their own national circumstances and priorities (UNGA 2015) in defining SDG targets and, hence, one-size-fits-all interventions do not exist.

13.6 Conclusion

The Sustainable Development Goals are highly ambitious and were adopted to stimulate action over the next 15 years in areas of critical importance for humanity and the planet (UNGA 2015). Digital Earth has untapped potential to improve the means of implementing the SDGs at both national and global scales. Through an extensive review of the recent literature and a case study of the operationalization of the SDG Indicator Framework in Australia, this chapter discussed information needs and promising operational initiatives underpinned by big Earth data and analytics, and, as importantly, multi-stakeholder partnerships. Digital Earth Australia is an example of the potential of Digital Earth to be an agent of 'partnerships for the goals', which can increase the availability of high-quality, timely and reliable data that is relevant in national contexts (SDG 17.18), and enhance regional and international cooperation on, and access to, science, technology and innovation (SDG17.6).

Acknowledgements The authors thank Geoscience Australia (Digital Earth Australia) and the Terrestrial Environment Research Network (TERN) for supporting the research. Prof. Richard Lucas also thanks the European Research Development Fund (ERDF) Ser Cymru Program and Welsh Government through the Living Wales project. Kate Owers (UNSW) and Sebastien Chognard (Aberystwyth University; Living Wales) are also thanked for their assistance in the development of the conceptual framework.

References

Allen C, Metternicht G, Wiedmann T (2017a) An iterative framework for national scenario modelling for the Sustainable Development Goals (SDGs). Sustain Dev 25(5):372–385

Allen C, Nejdawi R, El-Baba J et al (2017b) Indicator-based assessments of progress towards the Sustainable Development Goals (SDGs): a case study from the Arab Region. Sustain Sci 12(6):975–989

Allen C, Metternicht G, Wiedmann T (2018) Initial progress in implementing the Sustainable Development Goals (SDGs): a review of evidence from countries. Sustain Sci 13(5):1453–1467

Allen C, Metternicht G, Wiedmann T (2019) Prioritising SDG targets: assessing baselines, gaps and interlinkages. Sustain Sci 14(2):421–438

Asbridge E, Lucas R, Ticehurst C et al (2016) Mangrove response to environmental change in Australia's Gulf of Carpentaria. Ecology and Evolution 6: 3523–3539.

Asbridge E, Lucas R, Rogers K et al (2018). The extent of mangrove change and potential for recovery following severe Tropical Cyclone Yasi, Hinchinbrook Island, Queensland, Australia. Ecology and evolution. (21):10416–10434.

Australian Government (2018) Report on the Implementation of the Sustainable Development Goals (Voluntary National Review) 2018. https://dfat.gov.au/aid/topics/development-issues/2030-agenda/Pages/sustainable-development-goals.aspx. Accessed 19 Jun 2019

CEOS-GEO EO4SDG (2017) Earth Observations in support of the 2030 Agenda for Sustainable Development. https://www.earthobservations.org/documents/publications/201703_geo_eo_for_2030_agenda.pdf. Accessed 19 Jun 2019

Corbane C, Pesaresi M, Politis P et al (2017) Big Earth Data Analytics on Sentinel-1 and Landsat Imagery in Support to Global Human Settlements Mapping. Big Earth Data 1(1–2):118–144

Dhu T, Dunn B, Lewis B, et al (2017) Digital Earth Australia–unlocking new value from Earth Observation data. Big Earth Data. 1:64–74.

Di Gregorio A (2016) Land cover classification system: classification concepts. Software version 3. Land and Water Div. eng 187275. Italy, FAO

Dong J, Metternicht G, Hostert P et al (2019) Remote sensing and geospatial technologies in support of a normative land system science: status and prospects. Curr Opin Environ Sustain 38:44–52

Duke NC, Kovacs JM, Griffiths AD et al (2017) Large-scale dieback of mangroves in Australia's Gulf of Carpentaria: a severe ecosystem response, coincidental with an unusually extreme weather event. Mar Freshw Res 68(10):1816–1829

Foody GM, Ling F, Boyd DS et al (2019) Earth observation and machine learning to meet sustainable development goal 8.7: mapping sites associated with slavery from space. Remote Sens 11(3):266

Freire S, Schiavina M, Florczyk AJ et al (2018) Enhanced data and methods for improving open and free global population grids: putting 'leaving no one behind' into practice. Int J Digit Earth. https://doi.org/10.1080/17538947.2018.1548656

Gill T, Johansen K, Phinn S et al (2017) A method for mapping Australian woody vegetation cover by linking continental-scale field data and long-term Landsat time series. Int J Remote Sens 38(3):679–705

Giuliani G, Chatenoux B, De Bono A et al (2017) Building an earth observations data cube: lessons learned from the Swiss Data Cube (SDC) on generating Analysis Ready Data (ARD). Big Earth Data 1(1–2):100–117

Guo H (2017) Big data drives the development of Earth science. Big Earth Data 1(1–2):1–3

Hammer CL, Kostroch DC, Quiros G (2017) Big data: potential, challenges, and statistical implications. IMF staff discussion note, SDN/17/06 https://www.imf.org/en/Publications/Staff-Discussion-Notes/Issues/2017/09/13/Big-Data-Potential-Challenges-and-Statistical-Implications-45106. Accessed 19 Jun 2019

Hernandez M (2017) A Digital Earth platform for sustainability. International Journal of Digital Earth. 10(4):342–355

ICSU, ISSU (2015) Review of the sustainable development goals: the science perspective. International Council for Science (ICSU), Paris https://council.science/cms/2017/05/SDG-Report.pdf Accessed 19 Jun 2019

Inter-Agency and Expert Group on Sustainable Development Goals (2019) Tier classification for global SDG indicators. https://unstats.un.org/sdgs/iaeg-sdgs/. Accessed 19 Jun 2019

Kussul N, Lavreniuk M, Kolotii A et al (2019) A workflow for sustainable development goals indicators assessment based on high-resolution satellite data. Int J Digital Earth 1–13. https://doi.org/10.1080/17538947.2019.1610807

Le Blanc D (2015) Towards integration at last? The sustainable development goals as a network of targets. Sustain Dev 23(3):176–187

Lehmann A, Nativi S, Mazzetti P et al (2019) GEOEssential – mainstreaming workflows from data sources to environment policy indicators with essential variables. Int J Digital Earth 1–17. https://doi.org/10.1080/17538947.2019.1585977

Lewis A, Lymburner L, Purss MBJ et al (2016) Rapid, high-resolution detection of environmental change over continental scales from satellite data – the Earth Observation Data Cube. Int J Digital Earth 9(1):106–111

Lucas R, Mitchell A (2017) Integrated land cover and change classifications. In: Díaz-Delgado R, Lucas R, Hurford C (eds) The roles of remote sensing in nature conservation: a practical guide and case studies. Springer International Publishing, Cham, pp 295–308

Lucas R, Finlayson CM, Bartolo R et al (2018) Historical perspectives on the mangroves of Kakadu National Park. Marine and Freshwater Research.69 (7):1047–1063.

Lucas R, Mueller N, Siggins A et al (2019a) Land cover mapping using Digital Earth Australia. Data (in review)

Lucas R, Mueller N, Tissott B, et al (2019b) Land cover and evidence-based change mapping with the Digital Earth Australia's open source platform. ESA Living Planet Symposium, Milan, Italy, 13–17 May 2019. https://lps19.esa.int/NikalWebsitePortal/living-planet-symposium-2019/lps19/Agenda/AgendaItemDetail?id=29dd17de-09c7-419d-8558-6040526f6bcf. Accessed 14 Aug 2019

Lusseau D, Mancini F (2019) Income-based variation in sustainable development goal interaction networks. Nat Sustain 2(3):242–247

Lymburner L, Bunting P, Lucas R et al (2019) Mapping the multi-decadal mangrove dynamics of the Australian coastline. Remote Sens Environ 111185. https://doi.org/10.1016/j.rse.2019.05.004.

MacFeely S (2019) The Big (data) Bang: opportunities and challenges for compiling SDG indicators. Glob Policy 10(S1):121–133

Masó J, Serral I, Domingo-Marimon C et al (2019) Earth observations for sustainable development goals monitoring based on essential variables and driver-pressure-state-impact-response indicators. Int J Digital Earth 1–19. https://doi.org/10.1080/17538947.2019.1576787

Metternicht G, Lucas R, Bunting P et al (2018) Addressing mangrove protection in Australia: the contribution of earth observation technologies. In: IGARSS 2018–2018 IEEE international geoscience and remote sensing symposium, IEEE, pp 6548–6551

Metternicht G, Pagett M, Held A et al (2019) Big Earth Data enabling baseline data collection in support of SDG indicators: the experience of TERN Landscapes of Australia. In: Presentation 4th open science meeting of the global land programme, Bern, Switzerland, 24–26 April 2019

Mueller N, Lewis A, Roberts D et al (2016) Water observations from space: mapping surface water from 25 years of Landsat imagery across Australia. Remote Sens Environ 174:341–352

Nativi S, Santoro M, Giuliani G et al (2019) Towards a knowledge base to support global change policy goals. Int J Digital Earth 1–29. https://doi.org/10.1080/17538947.2018.1559367

Nilsson M, Griggs D, Visbeck M (2016) Policy: map the interactions between sustainable development goals. Nature 534(7607):320–322

Paganini M, Petiteville I, Ward S et al (2018) Satellite earth observations in support of the sustainable development goals. http://eohandbook.com/sdg/files/CEOS_EOHB_2018_SDG.pdf. Accessed 19 Jun 2019

Plag H-P, Jules-Plag S-A (2019) A goal-based approach to the identification of essential transformation variables in support of the implementation of the 2030 agenda for sustainable development. Int J Digital Earth 1–22. https://doi.org/10.1080/17538947.2018.1561761
Reyers B, Stafford-Smith M, Erb K-H et al (2017) Essential variables help to focus sustainable development goals monitoring. Curr Opin Environ Sustain 26–27:97–105
Roberts D, Dunn B, Mueller N (2018) Open data cube products using high-dimensional statistics of time series. In: IGARSS 2018 – 2018 IEEE International Geoscience and Remote Sensing Symposium, pp 8647–8650. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8518312. Accessed 16 Jun 2019
Sagar S, Roberts D, Bala B et al (2017) Extracting the intertidal extent and topography of the Australian coastline from a 28 year time series of Landsat observations. Remote Sens Environ 195:153–169
Sims N, Green C, Newnham G et al (2017) Good practice guidance. SDG Indicator 15.3.1: Proportion of land that is degraded over total land area. Version 1.0. https://www.unccd.int/sites/default/files/relevant-links/2017-10/Good%20Practice%20Guidance_SDG%20Indicator%2015.3.1_Version%201.0.pdf. Accessed 16 Jun 2019.
Soulard F, Grenier M (2018) Earth observation for official statistics at statistics Canada. Paper presented at the Environmental-economic accounts with Earth observation data workshop, ANU, Canberra, May 2018
Spangenberg JH (2017) Hot air or comprehensive progress? A critical assessment of the SDGs. Sustain Dev 25(4):311–321
Sudmanns M, Tiede D, Lang S et al (2019) Big Earth data: disruptive changes in Earth observation data management and analysis? Int J Digital Earth 1–19. https://doi.org/10.1080/17538947.2019.1585976
TERN (2017) TERN landscapes. https://www.tern.org.au/TERN-Landscapes-pg32473.html. Accessed 16 Jun 2019
UN Statistical Commission (2016) 47th Session of the UN Statistical Commission. https://unstats.un.org/unsd/statcom/47th-session/. Accessed 16 Jun 2019
United Nations Economic Commission for Europe (2016) Outcomes of the UNECE Project on using big data for official statistics. https://statswiki.unece.org/display/bigdata/Big+Data+in+Official+Statistics. Accessed 19 Jun 2019
United Nations General Assembly (2015) Transforming our World: the 2030 agenda for sustainable development. Resolution adopted by the general assembly on 25 September 2015: 70/1. http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E. Accessed 19 Jun 2019
Wunder S, Kaphengst T, Frelih-Larsen A (2018) Implementing land degradation neutrality (SDG 15.3) at national level: general approach, indicator selection and experiences from Germany. In: Ginzky H, Dooley E, Heuser IL et al (eds) International yearbook of soil law and policy 2017. Springer International Publishing, Cham, pp 191–219

Graciela Metternicht is a Professor in Environmental Geography at the University of New South Wales Sydney. Her research focuses on applications of geospatial technologies for environmental management and sustainability. Past appointments include Regional Coordinator of Early Warning and Assessment at UN Environment.

Norman Mueller is a Director of Product Development for Digital Earth Australia where he leads a team building time series tools and products for the monitoring and management of the environment from Earth observation data.

Richard Lucas holds a Sêr Cymru Research Chair at Aberystwyth University where his research focuses on the use of Earth observation data to better support monitoring, conservation, sustainable use and restoration of global environments.

Chapter 14
Digital Earth for Climate Change Research

Gensuo Jia, Li Zhang, Lanwei Zhu, Ronghan Xu, Dong Liang, Xiyan Xu and Tao Bao

Abstract Our planet is undergoing one of the most rapid climate changes in Earth's history. The current change is particularly significant because it is most likely a consequence of human activities since the 19th century. The Digital Earth platform, which includes Earth-orbiting satellites, ground-based observations, and other technologies for collecting, analyzing and visualizing data, has enabled scientists to see our climate and its impacts at regional and global scales. The Digital Earth platform offers valuable information on the atmosphere, biosphere, hydrosphere and cryosphere to understand Earth's past and present, and it supports Earth system models for climate prediction and projection. This chapter gives an overview of the advances in climate change studies based on Digital Earth and provides case studies that utilize Digital Earth in climate change research, such as in the observation of sensitive factors for climate change, global environmental change information and simulation systems, and synchronous satellite-aerial-ground observation experiments, which provide extremely large and abundant datasets. The mapping of climate extremes and impacts improves preparedness for climate change-related risks and provides robust evidence to support climate risk management and climate change adaptation for the public, decision makers, investors, and vulnerable communities. However, Digital Earth faces the challenges of multisource data coordination and integration, requiring international partnerships between governments and other organizations to advance open data policies and practices.

Keywords Digital Earth platform · Climate change · Sensitive variables · Information and model systems · Greenhouse gases

G. Jia (✉) · X. Xu · T. Bao
Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
e-mail: jiong@tea.ac.cn

L. Zhang · L. Zhu · D. Liang
Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing, China

R. Xu
National Satellite Meteorological Center, Beijing, China

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_14

14.1 Introduction

Global climate change has long been recognized as the most critical issue of the 21st century. The 2016 Paris Agreement within the United Nations Framework Convention on Climate Change (UNFCCC) highlights the importance and urgency of climate action. Climate-related changes are becoming evident at various spatial and temporal scales, accompanied by a record increase in the frequency of extreme climate events and the emergence of complex environmental issues. As a result, vulnerability to climate change is expected to expand spatially, threatening larger human populations as warming continues. Understanding climate change and delivering climate information with high precision has therefore become increasingly important, especially to assist governments and decision makers in implementing appropriate mitigation and adaptation policies.

The Earth system is a complex collection of interlinked subsystems that require multidimensional, multiscale and multitemporal datasets. Understandably, challenges and uncertainties in studying climate change and its impacts are largely due to the massive amount of data that is required, and the complexity of analyses that can translate data into useful information. Earth observation for this purpose has become an invaluable resource. Earth observations, during most of the history of science, have predominantly been recorded at the ground level with limited spatial coverage. Methods such as those developed by World Weather Watch in 1963 combined a series of single surface pictures to provide global coverage but lacked network density and vertical resolution. Geophysical and biological phenomena have also been generally insufficiently sampled. However, the growing diversity and improvement of sensors and sensing platforms has greatly diversified data sources, benefiting global climate change research in the past few decades through technologies that can increasingly provide a more accurate and precise picture of biological, physical, and chemical phenomena (Table 14.1). Moreover, satellite platforms and the development of UAVs and other technologies have made multitemporal observations feasible, which have allowed for investigations into large-scale processes that were traditionally not possible. Synoptic Earth observations have changed the way we understand the planet, from the first weather satellite that revealed astonishing cloud features to their utility to verify and improve our understanding of the coupling between the El Niño-Southern Oscillation and ocean currents. They have been used to study temperatures at various altitudes, atmospheric processes, the effects of snow on water circulation, the effects of global and regional factors on sea level changes, and other phenomena. From 1960 to 2011, 514 Earth observation satellites were launched worldwide, and 200 more launches are planned by 2030 (Guo 2014). The huge amount of data collected over the years provides a rich resource of information for climate change research. However, this big data presents challenges in data collection, characterization and analysis. Therefore, there is urgent need for a Digital Earth platform that can integrate multisource spatial information into a single platform and allow for integrated investigation into Earth observation data to generate climate change information.

Table 14.1 Summary of the functions of satellites related to global change research

Satellite	Function
TIROS series, Nimbus 4 and 7, ERS-1, ERS-2, Envisat	Monitors global stratospheric ozone depletion (including Antarctica and the Arctic)
Nimbus 7, ERS-2, Envisat, Aqua, Aura, MetOp	Detects tropospheric ozone
Explorer 7, TIROS, Nimbus	Measures radiation balance
TIROS series, ATS, SMS, MetOp	Produces weather images
Meteorological satellites, including the TIROS series, GOES and POES (NOAA), MetOp (Eumetsat), ERS-1, ERS-2, Envisat	Weather forecasting
Radarsat, Landsat, Aura, Terra, Jason, ERS-1, ERS-2, Envisat	Investigates ice flows in Antarctica and Greenland
Topex/Poseidon, ERS-1, ERS-2, Envisat	Detects mid-scale sea surface topography and important variables in ocean mixtures
TIROS-N and NOAA series, ERS-1, ERS-2, Envisat	Observations of oceanic contributions to climate change
Landsat, SPOT series	Agricultural land monitoring
LAGEOS, GPS	Confirms high-precision terrestrial reference frames
GCOM	Observes Earth water and carbon dioxide
TANSAT	Monitors atmospheric carbon dioxide concentration
FY	Used in weather forecasting, climate prediction, natural disaster and environmental monitoring, and resource development

Sources NRC (2008), Guo et al. (2015)

14.2 Digital Earth and the Essential Climate Variables

Climate change is highly heterogeneous over the globe, with strong regionality. The UNFCCC provides 34 essential climate variables (ECVs) that require contributions from Earth observations from space (Table 14.2) (Guo et al. 2015). The spatial attributes of ECVs make it possible to effectively observe them through space technology (Guo et al. 2014a), and the Digital Earth platform based on space technology plays an essential role in better understanding the spatial and temporal changes in the climate.

Table 14.2 Essential climate variables (ECVs) that are feasible for global implementation and have a high impact on UNFCCC requirements

Domain		Essential climate variables
Atmospheric	Surface	Air temperature, wind speed and direction, water vapor, pressure, precipitation, surface radiation budget
	Upper-air	Temperature, wind speed and direction, water vapor, cloud properties, Earth radiation budget (including solar irradiance)
	Composition	Carbon dioxide, methane, and other long-lived greenhouse gases; ozone and aerosols, supported by their precursors
Oceanic	Surface	Sea surface temperature, sea surface salinity, sea level, sea state, sea ice, surface current, ocean color (for biological activity), carbon dioxide partial pressure, ocean acidity
	Subsurface	Temperature, salinity, current, nutrients, carbon dioxide partial pressure, ocean acidity, oxygen, tracers, phytoplankton, marine biodiversity, and habitat properties
Terrestrial		River discharge, water use, groundwater, lakes, snow cover, glaciers and ice caps, ice sheets, permafrost, albedo, land cover (including vegetation type), fraction of absorbed photosynthetically active radiation (FAPAR, leaf area index (LAI), above-ground biomass, soil carbon, fire disturbance, soil moisture, terrestrial biodiversity, and habitat properties

Sources CEOS (2006), Guo et al. 2015

14.2.1 Earth Observation Data Parameters and Their Capabilities

Ground-based Earth observation systems such as rain gauge networks and radar have always been a major means of observing atmospheric structures and they are still being operated and maintained. However, satellite platforms have added valuable scientific data to monitor clouds, water vapor, precipitation, and wind at multiple spatial and temporal scales. Sensors such as the Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS) developed by the U.S., the Medium Resolution Imaging Spectrometer (MERIS) from the European Space Agency (ESA), and the international A-Train satellite systems have provided a wealth of information on clouds, rain, and pollutants, leading to a greater understanding of cloud pollution influences (Guo et al. 2015).

The cryosphere, consisting of lakes, river ice, snow cover, glaciers, ice caps, and frozen ground (including permafrost), is one of the most important parts of the climate system. Thus, changes in the cryosphere as well as in soil moisture and salinity are very important for monitoring global climate change, managing regional water resources, and investigating water and land ecosystems and global sea levels. Data from polar-orbiting and geostationary satellites (carrying visible/near-infrared sensors), such as the Geostationary Operational Environmental Satellite System (GOES), Landsat, MODIS, MERIS, and AVHRR, have been used to monitor the melt flow from snow cover and glaciers. This information is also valuable for the management of water resources and disasters, and has been utilized for flood disaster prediction and reservoir operation. Data from the Sea Winds scatterometers onboard the QuikSCAT satellites can monitor seasonal changes in ice, track giant icebergs, and provide daily maps of ocean ice at a 6-km resolution.

Earth observation satellites also provide hundreds of data products (Table 14.3) to monitor water quality, water color (e.g., chlorophyll, suspended solids, and turbidity) and sea surface temperatures. For example, the AVHRR, AATSR, and MODIS sensors provide data on sea surface temperatures (CEOS 2006; Guo et al. 2015). In addition, many satellites can obtain data on elevation measurements, geopotential heights, and terrain. For example, P-band synthetic aperture radar (SAR) can penetrate cloud cover and the vegetation canopy and is useful in tropical and northern forest research at high altitudes. Improved SAR such as advanced synthetic aperture radar (ASAR) and phased array L-band SAR (PALSAR) are available for agriculture, forestry, land cover classification, hydrology, and cartography.

The main characteristics of climate change are the trends in temperature, precipitation, polar ice cover, and sea level. A new generation of satellite systems and advanced sensors such as Suomi NPP, GPM, OSTM/Jason-2, ICESat-2, and SWOT

Table 14.3 Remotely sensed oceanographic parameters, their observational category, and representative sensors

Parameter	Observational category	Satellite/Sensor
Bio-optical	Visible to near-infrared	ENVISAT/MERIS, AQUA/MODIS, OrbView-2/SeaWiFS
Bathymetry	Visible to near-infrared	Landsat, SPOT, IKONOS
Sea surface temperature	Thermal infrared microwave radiometers	POES/AVHRR, GOES/Imager DMSP/SSM/I, TRMM/TMI
Sea surface roughness, wind velocities, waves and tides	Microwave scatterometers and altimeters Synthetic aperture radar	ERS-1 &-2/AMI QuikSCAT, RADARSAT-1
Sea surface height and wind speeds	Altimeters	Topex/Poseidon, Jason-1
Sea ice	Visible to near-infrared microwave radiometers, scatterometers and altimeters Synthetic aperture radar	POES/AVHRR DMSP/SSM/I ERS-1 &-2/AMI RADARSAT-1
Surface currents, fronts, and circulation	Visible to near-infrared, thermal infrared microwave scatterometers and altimeters	POES/AVHRR, GOES/Imager Topex/Poseidon, Jason-1
Surface objects-ships, wakes, and flotsam	Synthetic aperture radar	RADARSAT-1, ENVISAT/ASAR

Source Brown et al. (2007), Guo et al. 2015 have further improved our capability for space-based observation of these key parameters related to climate change. In addition to the space-based data, in situ data from ground measurements and reanalysis data are used to provide information on key indicators of climate change. The Copernicus Climate Change Service (C3S) compiles all the information obtained by the Copernicus environmental satellites, air and ground stations, and sensors to provide comprehensive pictures of the past, present, and future climate of Earth.

14.2.2 Heterogeneous Changes in Temperature

Heatwaves and rising temperatures have gained prominence in the context of global warming. Digital Earth technology is relatively mature for monitoring global land and sea surface temperatures, although the algorithms and retrieval accuracy need to be further improved, and satellite LST measurements have uncertainties caused by data accuracy and inconsistencies between sensors. Nevertheless, satellite measurements have been very useful in monitoring surface temperatures and detecting extreme temperature events.

Thermal infrared surface temperatures from satellite platforms are frequently integrated into data assimilation systems and reanalysis data systems for climate parameters, including NCEP/NCAR and NCEP/DOE, ERA-40, and JRA-25, which effectively improves the accuracy and reliability of datasets. The most widely used global land surface temperature datasets are the monthly data measured by the AVHRR thermal infrared band (4, 5) since 1982, the 8-day composite data derived from the MODIS thermal infrared band since 2000, the daily global LST and SST data provided by ENVISAT from the ESA, and the LST measured by Aster at small scales. The geostationary satellite system operated by the United States, Europe, China, Japan, and others provide low- and middle-latitude LST data at one-minute intervals. In addition, SeaWiFS, FY-2/4 and FY-3 can acquire LST data. The Suomi NPP satellite launched in 2012 carries a 750-meter spatial resolution Visible Infrared Imaging Radiometer Suite (VIIRS) sensor, and its surface temperature data quality was an improvement (Guillevic et al. 2014).

The monitoring and impact assessment of heat waves based on multisource thermal infrared remote sensing data have made important progress in recent years. Since the heat wave in central Europe in 2003, most large-scale heat wave events have been successfully monitored, including the large-area heat wave in southern Asia in the summer of 2010, the continued high-temperature anomaly in eastern Asia during the spring of 2013, the extreme low-temperature event that lasted several weeks in central and eastern North America in the winter of 2014, and the persistent heat wave that swept over southern Asia and western Europe in the spring and summer of 2015. Progressive improvement of Digital Earth's thermal environment platform that integrates multisensor and multiresolution spatial data can provide automatic and more accurate extreme temperature information, and support government decision making and public information services.

14.2.3 Heterogeneous Changes in Precipitation

The accuracy of precipitation estimation has improved over the years with satellite-based sensors. Satellite systems allow for continual monitoring and observation of precipitation on a global scale, which was only possible at fixed intervals with limited spatial coverage using conventional ground-based observation systems. Infrared sensors onboard geostationary satellites, passive microwave sensors carried by the polar-orbiting satellites, and active radar onboard the TRMM satellite and its successors have collected a huge wealth of data on precipitation over the years. The establishment of Global Precipitation Measurement (GPM) realized a satellite constellation with coordinated, seamless observation of global precipitation, indicating a new era of “digital precipitation”. GPM is an independent and complex project consisting of a core satellite and approximately eight other satellites. Its precipitation observation can reach a radius of 5 km, covering 90% of the global land and ocean surface at three-hour intervals, and can distinguish rainfall, snow, ice and other precipitation forms. It is much more advanced than the previous generation of TRMM.

Geostationary meteorological satellites such as FY-2, GOES, GMS, Meteosat, and MTSAT have seen improvements in multichannel scanning and real-time performance and have high spatial and temporal resolutions (from one-hour intervals to half-hour intensive observation, and 5-km and 1.25-km spatial resolution at nadir for the infrared and visible and near-infrared spectral channels, respectively). This makes them more effective in monitoring hazardous weather systems. Therefore, comprehensive application of multiple channels such as thermal infrared, visible light, near-infrared, and water vapor channels is an essential component of the Digital Earth platform for extreme precipitation monitoring.

14.2.4 Extreme Climate Events

Extreme climate events refer to serious deviation of the climate from its average state, including phenomena that are statistically less significant. Extreme climate events generally include high-temperature heat waves, extreme snow, strong tropical storms, floods, meteorological droughts, and natural fires. Space-based observation of extreme climate events consists of real-time warning and monitoring, rapid postdisaster assessment, and disaster risk reduction. This requires high spatial and temporal resolution satellite information and an efficient operational platform. This is both a challenge and an opportunity for Digital Earth. For instance, regarding disaster risk, the combination of multisource satellite data, land use data, and topographic data makes it possible to rapidly assess flood risk at the watershed and regional scales (Reager et al. 2014). Cold winter events have occurred frequently in Eurasia in the last 10 years, and extreme low temperatures have been record-breaking. Mori et al.

(2014) added Arctic sea ice data and SST to climate models and found that the reason for most cold winters is Arctic Oscillation (AO).

Digital Earth technology has shown great potential in disaster monitoring, emergency response, disaster assessment, and reconstruction. Disaster reduction is the most effective aspect of the Digital Earth platform, which can perform all-weather, all-day dynamic detection. Meteorological satellites, radar satellites, and high-resolution visible and near-infrared Earth observation satellites can be used to monitor rainfall, floods, and droughts in real time for emergency response. The Digital Earth platform can support rapid analysis of the statistics and distribution of flooded areas, flooded land use categories, and the number of people affected, especially when satellite data is combined with digital thematic maps such as administrative, land use, population, and socioeconomic maps.

14.3 Interactions Between Climate and Society Through Space and Time

14.3.1 Greenhouse Gas Exchange

The current global climate change is mainly attributed to rapidly increasing atmospheric concentrations of two greenhouse gases, carbon dioxide (CO_2) and methane (CH_4). Most of the body of research on greenhouse gases has focused on CO_2 rather than CH_4, which is a more potent greenhouse gas. The lack of high spatial and temporal resolution datasets on continuous flux is a major reason for the limited knowledge on CH_4 exchange (Holgerson and Raymond 2016). In the case of CO_2, the scientific community still lacks a detailed understanding. For example, according to existing ground measurements, 40–50% of the carbon dioxide produced by human activities remains in Earth's atmosphere, and the remaining 50–60% is considered to be absorbed by the ocean and ground vegetation. However, scientists do not know exactly where the carbon dioxide is stored, how this storage process occurs, and whether this process can limit the increase in carbon dioxide in the atmosphere. To date, spatiotemporal pattern studies of terrestrial carbon sources and carbon sinks based on space technology have been mainly achieved through satellite-based visible light and near-infrared band indexes. The 8-km inverted AVHRR continuous vegetation index data is the longest time series, since 1982, and the accuracy of the sixth generation of the MODIS (C6) vegetation index data has been greatly improved. In addition, Landsat, MERIS, VIIRS, SPOT Vegetation, and Sea-Viewing Wide Field-of-View Sensor (SeaWiFS) data are available.

A key parameter for monitoring the temporal and spatial patterns of terrestrial carbon sources and carbon sinks is the fraction of absorbed photosynthetically active radiation (FAPAR), which largely determines the total gross primary production (GPP) or carbon assimilation capacity. To date, more than six different global FAPAR spatial databases have been released, inverted from MODIS, MERIS, SeaWiFS, MODIS-TIP, SPOT-VEG, and AVHRR time series data; however, the data are highly uncertain. A systematic evaluation of more than 800 ground sample datasets revealed that they differed greatly between continents and biomes, and all were insensitive to vegetation coverage and needed further improvement (Pickett-Heaps et al. 2014). Chinese scholars have made costrengthening observations among 25 field flux stations and driven vegetation productivity models with flux data, satellite-based vegetation indexes, surface albedo, and soil moisture indexes, which have significantly improved the estimation of FAPAR and GPP on a regional scale (Wang et al. 2010). A recent improvement was the use of chlorophyll fluorescence data from the GOME satellite to drive vegetation productivity models and monitor global crop photosynthesis (Guanter et al. 2014).

The Orbital Carbon Observing Satellite (OCO-2) is a satellite launched by the United States in 2014 to monitor the global space-time distribution of carbon dioxide. It is mainly used to observe the carbon dioxide level of the Earth's atmosphere and to understand the role of humans in global climate changes caused by greenhouse gas emissions. The satellite carries a three-channel spectrometer for accurate measurements. OCO-2 collects approximately 8 million accurate global carbon dioxide measurements every 16 days, with a measurement accuracy of one in a million. With instruments such as spectrometers carried on satellites, scientists can dynamically measure carbon dioxide from different sources in the atmosphere and monitor the adsorption of carbon dioxide by oceans and forests. The acquisition of such a dynamic global carbon dioxide map will help reduce errors and improve the accuracy of forecasts for global warming.

Prior to this, in 2009, JAXA (Japan) launched GOSAT, the first satellite dedicated to detecting the concentration of greenhouse gases such as atmospheric CO_2. The satellite was equipped with high-precision observation equipment that used greenhouse gases such as carbon dioxide and methane to absorb infrared rays of a specific wavelength, and estimated the concentration of greenhouse gases by observing the infrared rays reflected from the surface. The goal of GOSAT was to observe the distribution of global CO_2 and CH_4, with a measurement precision of 2–3 ppm for CO_2 and approximately 15 ppb for CH_4, to capture the spatial variation in the carbon flux each year. GOSAT was equipped with thermal infrared and near-infrared sensors to obtain carbon observations as well as cloud and aerosol images. As the infrared rays passes through the atmosphere, a gas that forms a greenhouse effect, such as carbon dioxide, causes a specific wavelength to be absorbed, and the concentration of the gas can be calculated from these data. TanSat, launched by China, has further improved our ability to detect atmospheric CO_2 and other greenhouse gases.

Many countries including China are actively planning to launch satellites for the special detection of atmospheric CO_2 and other greenhouse gases. Integrating spatial data from these different sources with station observation data on the Digital Earth platform will greatly enhance the accuracy of detection and the technical support for climate change research.

14.3.2 Connectivity and Teleconnection in the Earth System

The Earth system as a whole, its components, and the various regional subsystems are connected and closely related. For example, in ocean-air interactions, the transfer of energy between the two is a teleconnection. We are gradually recognizing the importance of teleconnections in the climate system. For example, variability in the El Niño-Southern Oscillation (ENSO) model across the equatorial Pacific is linked to widespread distribution of floods, droughts, and forest fires in often arid or semiarid areas such as East Africa, tropical and subtropical Australia, and North America within the mid-latitudes and the western coast of South America. Another good example comes from Mori et al. (2014), who showed that most cold winters are attributed to AO changes caused by Arctic sea ice.

Studying and understanding teleconnections is an important challenge and an undertaking that can greatly benefit from utilization of the Digital Earth platform's capabilities of macroscopic multiparameter data integration to enable discovery of hidden and underlying connections in the Earth system and reveal the mechanisms to improve predictions of climate and weather. Various aspects remain to be identified and can benefit from the Digital Earth platform. For example, regarding ENSO and the North Atlantic Oscillation (NAO), we know relatively little about the teleconnection between the stratosphere and the Earth's surface. A strong vortex is formed over the polar region in winter, and the vortex intensity changes. When the vortex is strong, a tightly stable cycle is concentrated in the stratosphere, with little connection to the troposphere and the Earth's surface. When the vortex is weak, the control is not very stable, and it can generate a large-scale dynamic process. Therefore, it can be transmitted to the surface of the Earth through the convective top layer. It causes unusually cold weather at high latitudes, for example, in Scandinavia. When the Arctic vortex weakens, cold air flows outwards and downwards. Another example is the study by Zhang et al. (2019), which showed that the mean winter visibility throughout most of eastern China is negatively correlated with the preceding Antarctic Oscillation (AAO), especially in northern China. It emphasizes the important roles of sea surface temperature warming or cooling tendencies in the northwestern southern Indian Ocean (NSIO) and provides possible pathways through which NSIO warming may influence the atmosphere of northern China.

14.4 Impacts and Response

14.4.1 Ecosystems

Currently, the spatial data used to analyze the response of large-scale ecosystems to climate change are mainly acquired from long-term time series data from medium- and coarse-resolution optical satellite sensors such as AVHRR, MODIS, SPOT, VIIRS, SeaWiFS, and MERIS, which have inconsistencies between the sensors and their time spans (Guay et al. 2014). Several released global data series are generally based on the records of a single sensor. There are few data series from multisource data fusion and integration. However, satellite data often contain uncertainties caused by biases in different sensors and retrieval algorithms as well as inconsistencies between continuing satellite missions with the same sensor. Undetected drifts in sensor sensitivity have been cited as the main reason for the apparent spectrum of change. If the procedures for merging data from different time series are not well-developed and calibrated, the uncertainties can potentially be high in combined datasets. An integrated vegetation index dataset based on system calibration and data fusion is an important requirement for the Digital Earth platform.

Due to the complexity of ecosystem dynamics in the context of climate change, traditional methods based on single-satellite data have great limitations. By integrating and comparing multiple satellite datasets and ground observation data, the Digital Earth platform can dynamically and effectively display and analyze the trends of climate-related parameters.

14.4.2 Water Cycle and Water Resources

The global water cycle involves transformation, flow, and redistribution, and the redistribution of global and regional energy and regulation of the climate. The Earth observation system can quantitatively monitor many key parameters of the global water cycle, including various forms of precipitation (such as rainfall, hailstones, ice rain, and snow), atmospheric water vapor, surface evaporation, vegetation canopy transpiration, surface water, snow, continental glaciers, sea ice, soil moisture, and surface runoff.

Using the Digital Earth platform, global hydrology cycle models can be developed to reveal the controlling factors of terrestrial water cycling and trends in water resource patterns. It is expected to lead to a revolutionary solution to a series of key issues in Earth's multiple spheres of interactions from the perspective of Earth system dynamics, including global ocean-atmospheric interaction, land-atmospheric interaction and the boundary layer process, ocean-land correlation, and coastal ecosystem evolution.

14.4.3 Coastline, Urban Areas, and Infrastructure

Smajgl et al. (2015) employed remote sensing land use data, digital elevation data, and high-resolution climate models to simulate the scenario of a regional sea level rise of 30 cm by 2050. The study predicted that urban floods and sea water backflow would be severe downstream of the Greater Mekong Subregion and that the land use structure would change significantly.

The urban heat island effect accompanies the expansion of human settlements and is closely related to regional climate change. As the most active region of economic growth and urbanization, the urban heat island phenomenon in Asia, especially in China, has become an important issue in regional climate change. The Digital Earth system provides comprehensive spatial information about urban areas (Hu et al. 2015), human activity intensity (Zhou et al. 2014), and thermal infrared land surface temperature. It provides a scientific platform for research on urban heat islands at different spatial and temporal scales. Regarding the potential contributions of infrastructure to a warming climate, researchers have examined the impacts of urban expansion on the trends in air temperature by investigating the changes in urban land use around meteorological stations and analyzing the relationship between the rate of urban expansion and air temperature magnitudes (He et al. 2013). Urban heat islands can influence land-atmospheric energy exchange, the turbulence regime of atmospheric flow, and the microclimate, and can accordingly modify the boundary layer processes over urban canopy and downstream areas. Research showed that estimation of key urban morphology parameters using high- and medium-resolution satellite data and intense field measurement along urban-rural transects can improve the performance of regional climate models in capturing critical climate effects over large and rapidly expanding urban clusters (Jia et al. 2015; Feng et al. 2014; Wang et al. 2012).

14.5 Multisource Digital Earth for Studying Climate Change Phenomena

Earth is a large, complex system, broadly grouped into three subsystems: the atmosphere, oceans, and land surface. Climate change involves understanding changes in one of these subsystems and understanding how these systems interact, their impacts on one another, and the consequences of changes in any one of them or their subsystems. This requires rich scientific datasets quantifying sensitive climate factors, which is not possible without integration of data from multiple sources. These multisource datasets have been collected over the years through synchronous satellite-aerial-ground observation experiments (Fig. 14.1).

Multisource datasets allow for comprehensive, continuous, and diverse information on the Earth's surface. Similarly, multisensor remote sensing datasets enable dynamic (and in some cases real-time or near real-time) monitoring of Earth's systems. It has played a fundamental role in supporting modern data-driven scientific innovation. Effective use of multiplatform Earth observation data with multiple sensors helps avoid and mitigate issues related to information extraction and inversions that arise from the use of a single sensor.

These datasets have enabled researchers to explore new theories by developing new methodologies and assimilation models that can incorporate multisource/multisensor, heterogeneous spatial data to acquire precise information on

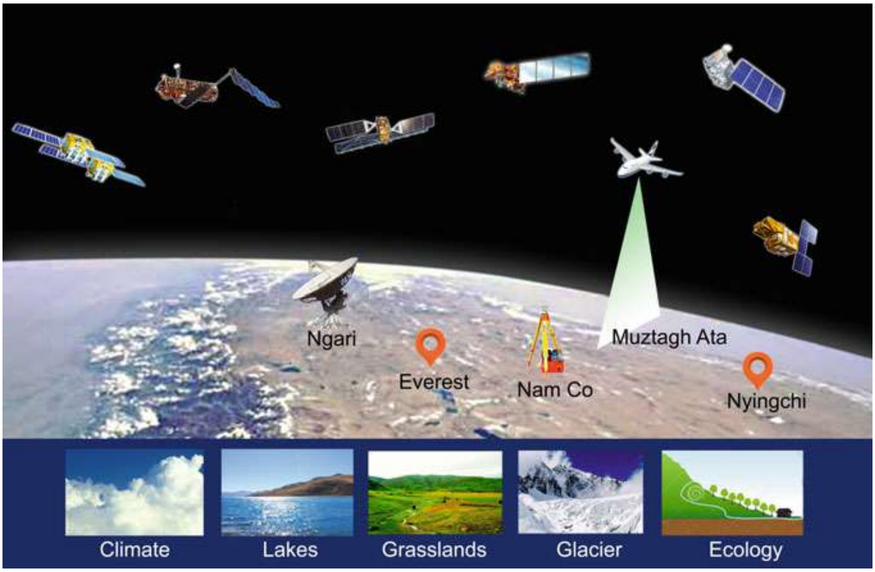


Fig. 14.1 Synchronous satellite-aerial-ground observation experiments on the Qinghai-Tibetan Plateau (revised from Guo et al. 2015)

sensitive climate factors and develop simulation platforms to understand regional climate change patterns. Multisensor Earth observations also provide long-term, stable spatial data for scientific research, compensating for uneven spatiotemporal observations, and play a fundamental supporting role in global change research.

The National Basic Research Program of China (973 Program) launched the project "Earth Observation for Sensitive Variables of Global Change: Mechanisms and Methodologies" on January 1, 2009. This was the first research project on Earth observation techniques for global change research in China. The project highlighted sensitive variables in terrestrial, oceanic and atmospheric systems based on big data from Earth observation from multiple platforms and multiband sensors, focusing on the development of new theories, technologies, and methods in these fields. The research scheme of the project is shown in Fig. 14.2. During the project, the new concept of moon-based Earth observation for global change monitoring was also widely discussed and considered as an efficient way to map the solid earth dynamics and radiation budget at the top of the atmosphere (Guo et al. 2014b, 2018).

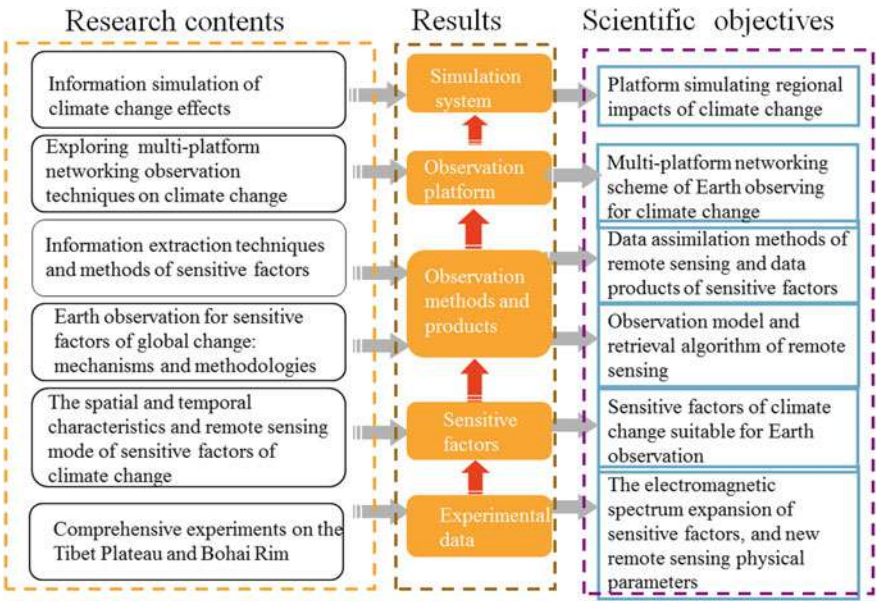


Fig. 14.2 Research scheme for the "Earth observation for sensitive variables of global change: mechanisms and methodologies" project (Guo et al. 2015)

14.5.1 Glaciers

Glaciers provide unique records and feedback that influence global climate change and are closely related to temperature, precipitation, and the material balance. The glaciers on the Tibetan Plateau have retreated considerably since the 1970s, and this rate of retreat has accelerated in recent years. In general, the retreat rate for glaciers covering less than 1 km^2 is faster than those of larger glaciers, but there are significant spatial differences. For example, glacial retreat was observed to be the fastest in the Himalayas and slower in the central plateau (Yao et al. 2003). It has been suggested that the retreat of the Himalayan glaciers is much more serious than expected (Ma et al. 2010). Consequently, with the rapid melting of glaciers, lakes supplied by the glacier melt water, such as Nam Co Lake (the highest lake on the central Tibetan Plateau), have expanded between 1976 and 2009 (Zhang et al. 2011; Guo et al. 2015).

A method for extracting glacier thickness has been developed based on interferometric synthetic aperture radar (InSAR) data and elevation data from the Geoscience Laser Altimeter System instrument aboard the Ice, Cloud, and land Elevation satellite (ICESat/GLAS14). As a result of calculations using the ICESat data along with the Shuttle Radar Topography Mission digital elevation model (SRTM DEM), a reduction of 0.63 m per year (water equivalent) was observed in the thickness of the Naimona'nyi glacier between 2000 and 2009 (Zong et al. 2013). This lies between the material balance of 0.56 m per year (water equivalent) and the glacier thickness reduction of 0.65 m per year (water equivalent) measured by GPS (Li et al. 2012). In general, glacial shrinkage decreases toward the interior plateau from the Himalayas, and the minimum degree of shrinkage occurs in the Pamir mountain range (Yao et al. 2012; Guo et al. 2015).

14.5.2 Lakes

Large fluctuations in lake surface area in a short time significantly influence water cycles and the local ecological environments. Studies have been conducted on lake areas, in addition to water level monitoring in different regions of the Tibetan Plateau using Landsat and ICESat data. Since 2003, a large spatial variation in lake area on the Tibetan Plateau has been observed, with a shrinkage of lakes in southern Tibet and an expansion trend for lakes in the Qiangtang region (Liao et al. 2013). In the Qaidam Basin, Qinghai Lake showed an expansion trend, and the annual rate of change in water volume in spring was greater than that in autumn. Gyaring Lake in the eastern Tibetan Plateau also showed an expansion trend that mirrored that of Qinghai Lake (Liao et al. 2013). Glacial melt is the dominant driver of the recent lake expansions on the Tibetan Plateau. By investigating detailed changes in the surface area and levels of lakes across the Tibetan Plateau from Landsat/ICESat data, Li et al. (2014) found a spatial pattern in the lake changes from 1970 to 2010 (especially after 2000). They observed a southwest-northeast transition from shrinking, to stable, to rapidly expanding lakes, which suggests a limited influence of glacial melt on lake dynamics. The plateau-wide pattern of lake area changes is related to precipitation variations and is consistent with the pattern of permafrost degradation induced by rising temperatures (Li et al. 2014; Guo et al. 2015).

14.5.3 Vegetation

The plant phenological period is closely related to climate change, and phenological changes influence the carbon balance of terrestrial ecosystems by affecting ecosystem productivity. The alpine vegetation on the Tibetan Plateau is extremely sensitive to global change. Zhang et al. (2013), Wang et al. (2015) used MODIS to analyze the response and driving factors of space observations of plant greenness and phenology (Fig. 14.3). Zhang et al. (2013) found that the normalized difference vegetation index (NDVI) showed a gradual increasing trend in the plateau during the growing seasons from 2000 to 2009. On the western Tibetan Plateau, the continuous decrease in precipitation resulted in a delay in the alpine grassland phenology; in the eastern part of the plateau, the precipitation continued to increase, resulting in an advance in the grassland phenology (Wang et al. 2015). In addition, Liu et al. (2014) found that the spring phenology of the grasslands on the Tibetan Plateau exhibited a stronger

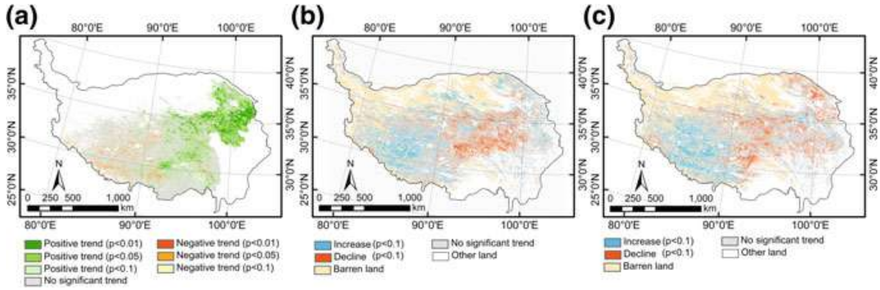


Fig. 14.3 Trends in **a** the growing season NDVI, **b** the start of the season, and **c** the end of the season on the Tibetan Plateau during 2000–2009 (Zhang et al. 2013; Wang et al. 2015; Guo et al. 2015)

response to changes in temperature at higher elevations than at lower elevations (Guo et al. 2015).

The remote sensing and monitoring of C_3 and C_4 grass species and their responses to climate change are mainly focused on the high-precision extraction of plant functional types and the transformation response of the grassland type to global climate change and human factors. In the U.S. Great Plains, vegetation with different functional types usually shows similar temporal trends in NDVI but different phenological characteristics (Wang et al. 2013). The onset of the growing season for C_3 grasses is earlier than that for C_4 grasses, and the growing season of C_3 grasses is longer. However, under mild weather conditions, C_3/C_4 short grasses have similar onsets of season dates and growing season lengths compared with C_3/C_4 tallgrasses (Wang et al. 2013). In northern China, a study by Guan et al. (2012) showed that temperate grassland was mainly occupied by C_3 species, yet C_4 species made an important contribution to the grassland biomass.

The fraction of photosynthetically active radiation (fPAR) is an important physiological parameter that reflects the growth of vegetation and is a key parameter for terrestrial ecosystem models and for reflecting global climate change (Fig. 14.4). Peng et al. (2012) found that the spatial variation in the global fPAR was affected

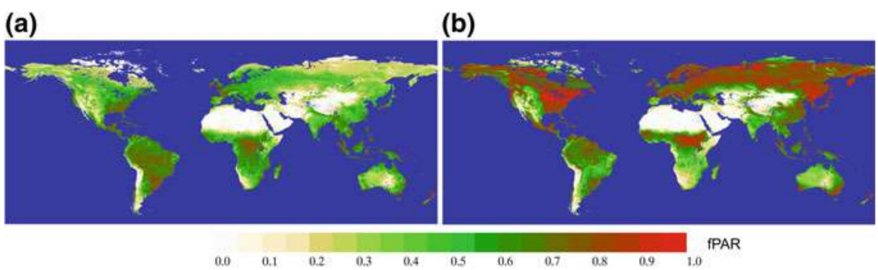


Fig. 14.4 Spatial patterns of global fPAR: **a** annual average fPAR in 2006, and **b** average fPAR in the latter half of August 2006. (Guo et al. 2015)

by the vegetation types as well as changes in the seasonal cycles. Temperature, precipitation and extreme drought have different effects on the fPAR. Climate change, deforestation, reforestation, and other human activities also significantly impact the fPAR in regions such as southeast Asia and the Three-North Shelter Forest area in China (Guo et al. 2015).

14.5.4 Radiation

(1) Impacts of aerosols on cloud cover and the regional radiation forcing effect

Based on satellite remote sensing data from aerosol-cloud-radiation and trace gases and meteorological observations, Xia (2010, 2012) analyzed long-term trends in the sunshine duration (SSD) and surface solar radiation and focused on the possible impacts of clouds on solar radiation in China over the last 50 years. The results indicated that the SSD and total cloud cover (TCC) showed a significant decreasing trend; however, with low-level cloud cover (LCC), a slight increasing trend was observed (Xia 2010). Short-term variability in the SSD is mainly determined by the amount of cloud cover, but the long-term change in the TCC cannot account for the decreasing trend in the SSD. Regarding the impacts of aerosols on clouds, Xia (2012) found that the data are inconsistent with the expectation that larger decreasing trends in cloud cover should be observed in regions with higher aerosol loading. Therefore, the aerosol effect on decreasing cloud cover in China does not appear to be supported by the results of their study (Guo et al. 2015).

(2) Spatiotemporal characteristics of land surface solar radiation in China

The land surface solar radiation in China and its temporal trends were calculated and the results demonstrate that previous studies overestimated the downward trend in land surface solar radiation in China (Tang et al. 2011). However, the aerosol abundance from human activities was still negligible on the Tibetan Plateau, and the decrease in solar radiation over the plateau was larger in magnitude than that for the rest of China after the 1970s. Further research revealed that the solar radiation on the Tibetan Plateau had continually decreased over the preceding 30 years due to the increasing water vapor and deep convective clouds. These increases were found to be connected to the warming climate and the enhanced effective convection energy of the Tibetan Plateau (Tang et al. 2011; Guo et al. 2015).

14.6 Digital Earth to Inform Climate Adaptation, Mitigation, and Sustainable Development

Effective strategies for climate change adaptation and mitigation require a comprehensive understanding of various underlying factors, including natural science, economics, society, and ethics. This makes climate change one of the most complex and challenging issues of modern times. Climate prediction and climate change projection are highly relevant to policy makers, investors, and vulnerable communities. The Digital Earth platform allows for investigations into many important processes that control the climate system, incorporates spatial dimensions at higher resolutions into the climate change context, and enables intuitive visual support for decisions and innovative actions. Strong visual and virtual demonstrations, supported by the Digital Earth platform, can help translate complex data into communicable information to support governments in decision and policy development and public information services.

Decades of Earth observation information is critical to improving predictions at different scales of climate projections. However, the existing remote sensing products have defects such as noise and time and space discontinuity (Brown et al. 2006; Jia et al. 2006). These defects severely constrain land surface processes and climate change simulations that are driven by spatial data parameters, and therefore reduce the reliability of climate change predictions and projections. It is necessary to synthesize multisensor remote sensing data to obtain high-quality and spatiotemporally continuous land surface observation data. The synthesis processes face the challenges of multisensor remote sensing data coordination and validation (Guo et al. 2015). These processes can greatly benefit from the Digital Earth data framework.

In addition to climate-sensitive environmental parameters, socioeconomic parameters characterize the demographic, socioeconomic, and technological driving forces underlying anthropogenic greenhouse gas emissions that have driven recent climate change and are key in the assessment of climate impacts, adaptation, and vulnerability. Conversely, the sensitivity, vulnerability, and adaptive capacity of socioeconomic systems also depend on their responses to climate change. The IPCC Technical Guidelines for Assessing Climate Change Impacts and Adaptations recommend the use of socioeconomic scenarios, with and without climate change, to assess climate impacts and adaptive responses. This adds a layer of complexity to predicting future scenarios and is only possible in the integrative environment provided by the Digital Earth platform. The challenges in implementing socioeconomic scenarios in Digital Earth include compatible scales that match the socioeconomic and satellite data, and rational assumptions that represent the evolution of key socioeconomic drivers.

The Digital Earth platform can also support the implementation of the UN Sustainable Development Goals (SDGs) by providing a conducive platform for information and data sharing, access, and use, and as a multisource data fusion platform. In the near future, Earth science will extensively make use of large amounts of data to monitor and predict continuously changing climatic environments. The Digital Earth platform can handle the challenge of geographical big data and the new emerging threats from climate change more systematically and specifically (Elder et al. 2016; Guo et al. 2017). This greatly enhances preparedness, rapid response, and adaptation to extreme events (such as extreme weather events) and facilitates understanding of the climate and projection of climate change.

In addition to geographical big data, a new form of geo-referenced data from the internet and social media, when combined with newly available observational, reanalysis, or other data sources on the Digital Earth platform, can potentially expand the scope of climate change studies greatly and increase the spatial and temporal scales addressed. For example, by using data from social networking sites, smart phones, and online experiments, we can assess the vulnerability to weather events and the impacts of local and national policies and programs in real time (Hernandez 2017).

Digital Earth has great potential for increasing our understanding of global climate change and its impacts on various dimensions. It is a powerful platform for policy support in climate change adaptation and mitigation. New developments in emerging technologies such as "big Earth data", citizen science, the blockchain, and artificial intelligence further enhance the power of Digital Earth to support studies and actions on climate change.

References

Brown ME, Pinzon JE, Didan K et al (2006) Evaluation of the consistency of long-term NDVI time series derived from AVHRR, SPOT-vegetation, SeaWiFS, MODIS, and landsat ETM + sensors. IEEE Trans Geosci Remote Sens 44(7):1787–1793

Brown CW, Connor LN, Lillibridge JL et al (2007) An Introduction to Satellite Sensors, Observations and Techniques. In: Miller R.L., Del Castillo C.E., Mckee B.A. (eds) Remote Sensing of Coastal Aquatic Environments. Remote Sensing and Digital Image Processing, vol 7. Springer, Dordrecht

CEOS (Committee on Earth Observation Satellites) (2006) Satellite observation of the climate system: the committee on earth observation satellites (CEOS) response to the global climate observing system (GCOS) implementation plan (IP). http://geodesy.unr.edu/hanspeterplag/library/earthobservations/CEOSResponse_1010A-1.pdf Accessed 15 March 2019

Elder M, Bengtsson M, Akenji L (2016) An optimistic analysis of the means of implementation for sustainable development goals: thinking about goals as means. Sustainability 8(9):962

Feng J, Wang J, Yan Z (2014) Impact of anthropogenic heat release on regional climate in three vast urban agglomerations in China. Adv Atmos Sci 31(2):363–373

Guan L, Liu L, Peng D (2012) Monitoring the distribution of C3 and C4 grasses in a temperate grassland in Northern China using moderate resolution imaging spectroradiometer normalized difference vegetation index trajectories. J Appl Remote Sens 6(1):063535

Guanter L, Zhang Y, Jung M et al (2014) Global and time-resolved monitoring of crop photosynthesis with chlorophyll fluorescence. Proc Natl Acad Sci USA 111(14):E1327–E1333

Guay KC, Beck PS, Berner LT et al (2014) Vegetation productivity patterns at high northern latitudes: a multi-sensor satellite data assessment. Glob Chang Biol 20(10):3147–3158

Guillevic PC, Biard JC, Hulley GC et al (2014) Validation of land surface temperature products derived from the Visible Infrared Imaging Radiometer Suite (VIIRS) using ground-based and heritage satellite measurements. Remote Sens Environ 154:19–37

Guo H (ed) (2014) Scientific satellites for global change research. Science Press, 978-7-03-039404-0

Guo H, Fu W, Li X et al (2014a) Research on global change scientific satellites. Sci China Earth Sci 57(2):204–215
Guo H, Ding Y, Liu G et al (2014b) Conceptual study of lunar-based SAR for global change monitoring. Sci China Earth Sci 57(8):1771–1779
Guo H, Zhang L, Zhu L (2015) Earth observation big data for climate change research. Adv Clim Chang Res 6(2):108–117
Guo H, Liu Z, Jiang H et al (2017) Big earth data: a new challenge and opportunity for digital earth's development. Int J Digit Earth 10(1):1–12
Guo H, Liu G, Ding Y (2018) Moon-based earth observation: scientific concept and potential applications. Int J Digit Earth 11(6):546–557
He Y, Jia G, Hu Y et al (2013) Detecting urban warming signals in climate records. Adv Atmos Sci 30(4):1143–1153
Hernandez M (2017) A digital earth platform for sustainability. Int J Digit Earth 10(4):342–355
Holgerson M, Raymond P (2016) Large contribution to inland water CO2 and CH4 emissions from very small ponds. Nat Geosci 9(3):222–226
Hu Y, Jia G, van Genderen J et al (2015) Assessing surface albedo change and its induced radiation budget under rapid urbanization with landsat and GLASS data. Theor Appl Climatol 123(3–4):711–722
Jia GJ, Epstein HE, Walker DA (2006) Spatial heterogeneity of tundra vegetation response to recent temperature changes. Glob Chang Biol 12(1):42–55
Jia G, Xu R, Hu Y et al (2015) Multi-scale remote sensing estimates of urban fractions and road widths for regional models. Clim Chang 129(3–4):543–554
Li Z, Xing Q, Liu S et al (2012) Monitoring thickness and volume changes of the dongkemadi ice field on the Qinghai-Tibetan Plateau (1969–2000) using shuttle radar topography mission and map data. Int J Digit Earth 5(6):516–532
Li Y, Liao J, Guo H et al (2014) Patterns and potential drivers of dramatic changes in Tibetan lakes, 1972–2010. PLoS ONE 9(11):e111890
Liao J, Shen G, Li Y (2013) Lake variations in response to climate change in the Tibetan Plateau in the past 40 years. Int J Digit Earth 6(6):534–549
Liu L, Liu L, Liang L et al (2014) Effects of elevation on spring phenological sensitivity to temperature in Tibetan Plateau grasslands. Chin Sci Bull 59(34):4856–4863
Ma L, Tian L, Pu J et al (2010) Recent area and ice volume change of kangwure glacier in the middle of Himalayas. Chin Sci Bull 55(20):2088–2096
Mori M, Watanabe M, Shiogama H et al (2014) Robust Arctic sea-ice influence on the frequent Eurasian cold winters in past decades. Nat Geosci 7(12):869–873
NRC (National Research Council) (2008) Earth Observations from Space: the First 50 Years of Scientific Achievements. The National Academies Press, Washington, D.C.
Peng D, Zhang B, Liu L et al (2012) Seasonal dynamic pattern analysis on global FPAR derived from AVHRR GIMMS NDVI. Int J Digit Earth 5(5):439–455
Pickett-Heaps CA, Canadell JG, Briggs PR et al (2014) Evaluation of six satellite-derived Fraction of Absorbed Photosynthetic Active Radiation (FAPAR) products across the Australian continent. Remote Sens Environ 140:241–256
Reager J, Thomas B, Famiglietti J (2014) River basin flood potential inferred using GRACE gravity observations at several months lead time. Nat Geosci 7(8):588–592
Smajgl A, Toan TQ, Nhan DK et al (2015) Responding to rising sea levels in the Mekong delta. Nat Clim Chang 5(2):167–174
Tang W-J, Yang K, Qin J et al (2011) Solar radiation trend across China in recent decades: a revisit with quality-controlled data. Atmos Chem Phys 11(1):393–406
Wang H, Jia G, Fu C et al (2010) Deriving maximal light use efficiency from coordinated flux measurements and satellite data for regional gross primary production modeling. Remote Sens Environ 114(10):2248–2258
Wang J, Feng J, Yan Z et al (2012) Nested high-resolution modeling of the impact of urbanization on regional climate in three vast urban agglomerations in China. J Geophys Res Atmos 117:D21103

Wang C, Hunt ER, Zhang L et al (2013) Phenology-assisted classification of C3 and C4 grasses in the US Great Plains and their climate dependency with MODIS time series. Remote Sens Environ 138:90–101

Wang C, Guo H, Zhang L et al (2015) Assessing phenological change and climatic control of alpine grasslands in the Tibetan Plateau with MODIS time series. Int J Biometeorol 59(1):11–23

Xia X (2010) Spatiotemporal changes in sunshine duration and cloud amount as well as their relationship in China during 1954–2005. J Geophys Res Atmos 115:D00K06

Xia X (2012) Significant decreasing cloud cover during 1954-2005 due to more clear-sky days and less overcast days in China and its relation to aerosol. Ann Geophys 30(3):573–582

Yao T, Wang Y, Liu S et al (2003) Recent glacial retreat in high Asia in China and its impact on water resource in Northwest China. Sci China Ser D Earth Sci 47(12):1065–1075

Yao T, Thompson L, Yang W et al (2012) Different glacier status with atmospheric circulations in Tibetan Plateau and surroundings. Nat Clim Chang 2(9):663–667

Zhang B, Wu Y, Zhu L et al (2011) Estimation and trend detection of water storage at Nam Co Lake, central Tibetan Plateau. J Hydrol 405(1–2):161–170

Zhang G, Zhang Y, Dong J et al (2013) Green-up dates in the Tibetan Plateau have continuously advanced from 1982 to 2011. Proc Natl Acad Sci U S A 110(11):4309–4314

Zhang Z, Gong D, Mao R et al (2019) Possible influence of the Antarctic oscillation on haze pollution in North China. J Geophys Res Atmos 124(3):1307–1321

Zhou Y, Smith SJ, Elvidge CD et al (2014) A cluster-based method to map urban area from DMSP/OLS nightlights. Remote Sens Environ 147:173–185

Zong J, Ye Q, Tian L (2013) Recent Naimona'Nyi Glacier surface elevation changes on the Tibetan Plateau based on ICESat/GLAS, SRTM DEM and GPS measurements. Chin Sci Bull 59(21):2108–2118

Gensuo Jia is a professor at CAS Institute of Atmospheric Physics and the Director of START Regional Center for East Asia. He serves as coordinating lead author of IPCC Special Report on Climate Change and Land (SRCCL). He has broad research interests in terrestrial ecology and atmospheric sciences, including remote sensing of land surface processes, ecosystem-climate interactions, climate extremes, and climate change impacts.

Li Zhang is a Professor in Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences. She researches remote sensing applications for land-surface dynamics and coastal environment. She serves as the Steering Committee Member of the GEO Blue Planet. She has published over 110 scientific papers.

Lanwei Zhu is an associate professor, Institute of Remote Sensing and Digital earth, Chinese Academy of Sciences. She researches nature heritage change, coral reefs change and bleaching monitoring based on remote sensing.

Ronghan Xu is an associate professor at the National Satellite Meteorological Center, China Meteorological Administration (CMA). Her researches focus on climate change and extremes, land use/cover change, urban heat island, and thermal infrared remote sensing. She served as a lead author of global environmental outlook youth edition.

Dong Liang received B.Sc. and M.Sc. degrees in applied mathematics from the University of Hull, U.K. (2006), and Mälardalen University, Sweden (2009). He is pursuing a Ph.D. in cartography and GIS at the University of Chinese Academy of Sciences, with a focus on Big Earth Data and polar remote sensing.

Xiyan Xu is an associate professor at the Institute of Atmospheric Physics, Chinese Academy of Sciences (CAS). Her research fields cover land-atmosphere interactions, climate change, and climate effects of land use/cover change. She serves as a chapter scientist of IPCC Special Report on Climate Change and Land (SRCCL).

Tao Bao is an assistant professor at the Institute of Atmospheric Physics, Chinese Academy of Sciences (CAS). He researches changing soil and vegetation processes under climate change, land-atmosphere carbon flux, and interactive changes of climate and ecosystems in polar regions using field and satellite observations.

Chapter 15
Digital Earth for Disaster Mitigation

Milan Konecny, Temenoujka Bandrova, Petr Kubicek, Silvia Marinova, Radim Stampach, Zdenek Stachon and Tomas Reznik

Abstract This chapter describes the state-of-the-art of the potential of Digital Earth for progressively better solutions for disaster mitigation. The chapter illustrates the use of strong Digital Earth tools for data sharing and important potential for users, such as 2D or multi-D visualizations. Milestones of developments in early warning, disaster risk management and disaster risk reduction concepts are highlighted as a continuous movement between sustainable development and original concepts of disaster risk reduction. Improved solutions have been based on new research directions formulated in Sustainable Development Goals tasks and by expanding the possibilities of new effective solutions via newly organized data ecosystems generated by the United Nations Global Geospatial Information Management, the Group on Earth Observations and the Group on Earth Observations System of Systems, Copernicus and, more recently, the Digital Belt and Road initiative. The new trends in spatial big data are emphasized; the most important for disaster risk reduction are the basic theses of the U.N. Conference in Sendai. This chapter describes three aspects: innovative Digital Earth development, national and local disaster risk assessment and the benefits arising from the use of maps and dynamic data, and analyses of the contributions of cartography to disaster risk reduction.

Keywords Digital earth potentials · Big data · Risk assessment · Risk mapping technology · U.N. GGIM · DBAR

15.1 Introduction

In this chapter, we describe the state-of-the-art potential of Digital Earth (DE) for progressively better solutions for disaster mitigation.

M. Konecny (✉) · P. Kubicek · R. Stampach · Z. Stachon · T. Reznik
Laboratory on Geoinformatics and Cartography, Department of Geography, Masaryk University, Brno, Czech Republic
e-mail: konecny@geogr.muni.cz

T. Bandrova · S. Marinova
Laboratory on Cartography, University of Architecture, Civil Engineering and Geodesy, Sofia, Bulgaria

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_15

For over 20 years, DE has witnessed an ebb and flow in interest from the world's scientific community. Initially, it sought a place between activities focused strictly on maps, data and information (Global Map—GM, Global Spatial Data Infrastructure—GSDI, etc.). Later, it began to push through with a comprehensive concept and an emphasis on the need to share and integrate data and information, and impetus and knowledge from the scientific realm, the private sector, and the needs of people in different parts of the Earth. Today, novel solutions are expected from DE, which will also significantly help realize disaster risk reduction (DRR) and Disaster Mitigation projects. Al Gore (former vice president of the USA) described a concept and definition of Digital Earth in his speech in Los Angeles on January 1998, saying it is: "A multiresolution, three-dimensional representation of the planet, into which we can embed vast quantities of geo-referenced data" (Gore 1998). In 2008, Goodchild noted that "Digital Earth includes four aspects: visualization, ease of use, interoperability and mashups, modelling and simulation" (Goodchild 2008). Some of the best analyses of the potential of the DE concept in the European Union (EU) are the SWOT analyses by De Longueville et al. (2010a, b). Studies showed positive and attractive aspects based on the political and economic support of influential countries such as the USA, China and, more recently, Russia. They also found obstacles originating from overly complex DE approaches that did not fit the research concepts of the EU. Clarification of DE leadership was also an issue. These aspects are all important for finding more successful approaches to solve disaster mitigation and DRR problems that are natural, societal or economical, as well as complex ones including known and unidentified factors. In addition, knowledge and new technologies are developing. We now have access to new near- to real-time information resources such as Prevention Web, the knowledge platform for disaster risk reduction managed by the U.N. Office for Disaster Risk Reduction (U.N. DRR), and research analyzing some of the unsuccessful efforts in developed countries such as those during Hurricane Katrina and recommending adequate steps in the future.

Section 15.2 describes the terminology used in disaster mitigation and this chapter and as well as some of the supportive efforts of international scientific organizations. Section 15.3 describes the development of early warning (EW), disaster risk management (DRM) and disaster risk reduction (DRR) concepts. Section 15.4 describes Digital Earth for the future of disaster mitigation and DRR and innovative support of the implementation of the Sendai Framework and existing geospatial projects, including the U.N. Global Geospatial Information Management (U.N. GGIM), Copernicus, Global Earth Observation System of Systems (GEOSS) and Digital Belt and Road (DBAR). Section 15.5 introduces national and local disaster risk assessment and the benefits arising from the usage of maps and dynamic data. Section 15.6 analyzes and shows the development of selected disaster risk mapping approaches and technologies with examples of adaptation principles, context map composition and existing symbol systems. Studies have attempted to recognize how users and inhabitants understand information from databases, maps and specialized models. The final Sect. 15.7 discusses expected developments in the research and technology background in the near-future. It will be necessary to accelerate the creation of new concepts from new knowledge (like from the Hyogo Framework) and new environments created by the realization of ideas of the U.N. GGIM and Chinese DBAR. All these approaches were developed on the same background as part of new data and information media, demonstrating how the potential is open to all of society as well as specialists and decision makers. Some of the approaches, such as mobile tools and digital maps, are described in this chapter.

15.2 Terminology and Research Organization Efforts

A very important aspect of new approaches is the terminology. The United Nations International Strategy for Disaster Reduction (U.N. ISDR) created the first terminology from the fields of early warning, disaster risk management and disaster reduction, which has been updated according to development the field. In this chapter, selected terminology from the U.N. ISDR is used.

The definitions of disaster mitigation, emergency, disaster damage, disaster impact, disaster management, emergency management, disaster risk, acceptable risk, residual risk, disaster risk assessment, disaster risk management, disaster risk reduction, early warning system, multi-hazard early warning system, and vulnerability can be found in the U.N. ISDR (2009).

There are two globally operating organizations, the U.N. ISDR and Integrated Research on Disaster Risk (IRDR), which formulate global tasks in the disaster risk reduction (DRR) area. There are also activities in important world organizations and by members of the International Science Council (ICSU). The first working group and later the Commission Cartography for Early Warning and Disaster Risk Management were founded within the International Cartographic Association—ICA (in 2004 and 2007, respectively, arranged by Konecny). The activities of the International Society for Photogrammetry and Remote Sensing (ISPRS), which started the GI4DM organization, were also very fruitful as well as those of the International Federation of Surveyors (FIG), which was organized during Working Week 2016 in Christchurch, New Zealand, at the Recovery from Disaster conference.

15.3 Development of Early Warning (EW), Disaster Risk Management (DRM) and Disaster Risk Reduction (DRR) Concepts

In the past, DRM was solved together with problems of the environment, subsequently developed relatively separately, and a new DRR trend enhanced their close cooperation in contemporary sustainable development efforts. There are two lines of development in U.N. documents in approaches to crisis situations, both natural and anthropogenic. They are:

(1) Environmental, linked to finding the most appropriate environmental approaches to solve planet Earth's problems. They are mainly oriented around concepts of sustainable development (SD). As a first important document mentioning natural disasters in the Report on Approaches to Crisis Management Issues Related to Development, U.N. environmental policies were created at the United Nations Conference on the Human Environment in Stockholm on 5–16 June 1972 (http://www.biblebelievers.org.au/gc1972.htm). Later, this approach was documented at the United Nations Conference in Rio de Janeiro in 1992, in Johannesburg in 2002 and at many others.

(2) Crisis risk management (early warning, disaster management and disaster risk reduction). The second line of development includes the Yokohama and Hyogo World Conferences (1994 and 2005), the Global Platform for Disaster Risk Reduction in Geneva in 2010 and the key concept of the "U.N. International Strategy for Disaster Reduction" (ISDR—United Nations International Strategy for Disaster Reduction). Another concept was developed in disaster risk research, which addresses the problem of natural and human-induced environmental hazards in IRDR (Integrated Research on Disaster Risk) (Konecny et al. 2010).

Three United Nations Conferences focused on DRR have been held. First, the World Disaster Reduction Conference in Yokohama in 1994, which defined the Yokohama Strategy and Plan of Action for a Safer World: guidelines for natural disaster prevention, preparedness and mitigation. The Second World Conference on Disaster Reduction was held in Kobe, Japan from 18 to 22 January, 2005. The Hyogo Framework for Action (2005–2015) (HFA): Building the Resilience of Nations and Communities to Disasters was an outcome of the 2005 conference. The HFA set five specific priorities for action: (1) making disaster risk reduction a priority; (2) improving risk information and early warning; (3) building a culture of safety and resilience; (4) reducing the risks in key sectors; and (5) strengthening preparedness for response (WCDRR 2016). The third conference was the Third U.N. World Conference on Disaster Risk Reduction in Sendai, Japan in 2015 (United Nations General Assembly 2015). The goals and role of research in the realization of these topics are described in Sect. 15.4 of this chapter. The Sendai Framework materials highlighted the need to tackle disaster risk reduction and climate change adaption when setting the Sustainable Development Goals, particularly in light of the insufficient focus on risk reduction and resilience in the original Millennium Development Goals (WCDRR 2016).

15.4 Digital Earth for the Future of Disaster Mitigation and DRR: Innovative Support of the Implementation of the Sendai Framework

15.4.1 Sendai Disaster Reduction Conference Targets

In the Third U.N. World Conference (U.N. DRR) on March 14, 2015 in Sendai, Japan, the Sendai Framework for Disaster Risk Reduction 2015–2030 was adopted (United Nations General Assembly 2015). The U.N. DRR conference is a culmination of contemporary state-of-the-art approaches to solve the problems of risks and disasters on our planet. The conference materials mentioned the role of Information and Communication Technologies (ICT), geographical information system (GIS), remote sensing, mapping, sensors, and volunteered geographic information. The document does not mention explicitly Digital Earth, but the proposed solutions follow lines defined by Digital Earth pioneers and updated according to research frontiers in the world. The necessity of design for deep integration of data and information and the necessity of offering products to specialists, customers and all society in an understandable way were emphasized.

The Sendai Framework defined four new priorities of action:

- Priority 1: Understand disaster risk;
- Priority 2: Strengthen disaster risk governance to manage disaster risk;
- Priority 3: Invest in disaster risk reduction for resilience;
- Priority 4: Enhance disaster preparedness for effective response and "Build Back Better" in recovery, rehabilitation and reconstruction (United Nations General Assembly 2015).

The priorities are equally important to find better solutions, and the Digital Earth concept should be useful in addressing all of them. We discuss the priority 1 intentions here. Researchers know enough about individual disasters, but are weak in their knowledge when disasters are combined, as in the Fukushima nuclear power station collapse or the Wenchuan earthquake. It is very valuable that solutions are being accepted at global, national, regional and local levels. In priority 1: Understanding disaster risk, on national and local levels, there are requests to develop, periodically update and disseminate location-based disaster risk information such as risk maps to decision makers, the general public and communities at risk of exposure to a disaster in an appropriate format by using applicable geospatial information technology. In addition, local and national organizations promote real-time access to reliable data, make use of space and in situ information, including geographic information systems (GIS), and use information and communication technologies innovations to enhance measurement tools and the collection, analysis and dissemination of data.

The DRR framework defined in Sendai is inextricably linked with the main U.N. document defining the Sustainable Development Goals 2015–2030 (SDGs).

15.4.2 Global Development Policy Framework (GDPF)

With other U.N. documents such as the Sendai Framework for DRR 2015–2030, the SIDS Modalities of Action (SAMOA) Pathway, the Addis Ababa Action Agenda, the Paris Agreement on Climate Change and the HABITAT III Urban Agenda, the SDGs created a newly formulated Global Development Policy Framework (GDPF) (Fig. 15.1).

In addition to natural disasters, there are new issues connected with problems of cities or megacities from the geospatial information perspective in particular and for DE in general. These problems are defined in another activity of the GDPF—HABITAT III. Its key document “The New Urban Agenda” was adopted at the United Nations Conference on Housing and Sustainable Urban Development (Habitat III) in Quito, Ecuador (United Nations 2016) and represents a shared vision for a better and more sustainable future. If well-planned and well-managed, urbanization can be a powerful tool for sustainable development for both developing and developed countries. The conference reached a critical point in understanding that cities can be the source of solutions to, rather than the cause of, the challenges that our world is facing today.

The New Urban Agenda presents a paradigm shift based on the science of cities; it lays out standards and principles for the planning, construction, development, management, and improvement of urban areas. The agenda also incorporates a new recognition of the correlation between good urbanization and development. The New Urban Agenda realizes the 2030 Agenda for Sustainable Development, especially

Fig. 15.1 Global development policy framework. *Source* UN-GGIM: strengthening the global data ecosystem, by Scott, ©2018 United Nations. Reprinted with the permission of the United Nations

Goal 11 on Sustainable cities and communities. It also planned to adopt and implement DRR and management, reduce vulnerability, build resilience and responsiveness to natural and human-made hazards and foster the mitigation of and adaptation to climate change. DRR is aimed at preventing new risk, reducing existing disaster risk and managing residual risk, all of which contribute to strengthening resilience and therefore to the achievement of sustainable development. DRR is the policy objective of disaster risk management, and its goals and objectives are defined in disaster risk reduction strategies and plans.

To improve the quality of solutions in disaster mitigation and DRR, U.N. member states should facilitate the strengthening and normative capacity-building of global geospatial information management in support of the implementation of the 2030 Agenda. Efforts include promoting the use of geospatial information systems and services for modern mapping, methodological development, national and regional capacity-building, setting of standards, data collection, dissemination and sharing, and better integration of geospatial and statistical information systems of U.N. Member States.

15.4.3 U.N. GGIM

A newly established Global Data Ecosystem by the U.N. Global Geospatial Information Management (U.N. GGIM) will support realization of the SDGs, including all aspects linked with DRR, to respond to global data ecosystem needs. It helps to develop the global understanding of geospatial information and, in a second step, its coordination, coherence and implementation. The vision is to position geospatial information to address global challenges and missions to ensure that geospatial information and resources are coordinated maintained, accessible, and used effectively and efficiently by member states and society to address key global challenges in a timely manner.

In the U.N. GGIM, Scott defined the data needs for the 2030 Agenda as follows (Scott 2018): “The scope of the 2030 Agenda requires high-quality and disaggregated data that are timely, open, accessible, understandable and easy to use for a large range of users, including for decision making at all levels. There is a need for a reporting system on the SDGs that would have benefit from the subnational (local) to the national level; and allow for global reporting that builds directly on the data shared by countries. It is important to create an opportunity for countries to directly contribute to the global reporting. While the challenges are immense, the digital technology that is available today allows the necessary transformation. An aspiration is to strengthen countries’ national geospatial and statistical information systems to facilitate and enable a ‘data ecosystem’ that leverages an accessible, integrative and interoperable local to global system-of-systems.”

The U.N. GGIM is the newest initiative to qualitatively improve the potential to solve the problems of the world, including disaster mitigation. In addition, other important initiatives have the same aim in specific regions of the World—e.g., Copernicus for Europe and the Digital Belt and Road (DBAR) initiative in Asia.

15.4.4 Copernicus—A European Contribution to GEOSS

Copernicus (formerly Global Monitoring for Environment and Security—GMES) is a European project based on data received from Earth observation satellites and ground-based information. These data are coordinated, analyzed and prepared for end users. Through Copernicus, the state of our environment and its short-, medium- and long-term evolution are monitored to support policy decisions and investments. Copernicus plays key role in EU EW, DRM and DRR efforts. Copernicus mainly supports decision making by institutional and private actors. Decisions can concern new regulations to preserve our environment or urgent measures in the case of natural or man–made catastrophes (i.e., floods, forest fires, water pollution) on a global scale. The services are used by environmental agencies, local, regional and national authorities, and civil protection organizations. The new observation techniques and analysis of data will allow for these actors to better anticipate potential threats, to intervene in a timely manner and to increase the efficiency of the intervention. Figure 15.2 shows the structure and purposes of Copernicus.

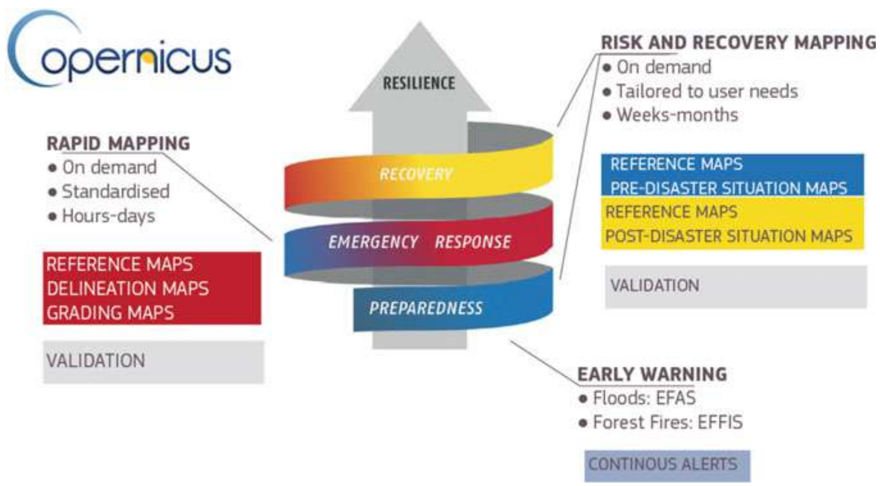


Fig. 15.2 Structure of Copernicus. Adapted from: EC (2019). Used with permission: Copernicus EU, European Commission

Copernicus (and its INSPIRE component) is the European contribution and participation in the worldwide monitoring and management of planet Earth organized by the Group on Earth Observation (GEO). The global community acts together for a synergy of all techniques of observation, detection and analysis. At the World Summit on Earth Observation in Washington in July 2003, the Group on Earth Observations (GEO) was established with the goal of addressing the information requirements for the environment on a global scale. In Brussels in February 2005, a 10-year implementation plan of an integrated global earth observation system of systems (GEOSS) was defined. A number of operational systems for supporting disaster response have made steady to strong progress. Collaborative supersites have been established for the scientific community to monitor and analyze volcanoes and earthquakes more rapidly and effectively; for example, supersites have improved the assessment of earthquakes in Haiti, China, Chile, and Indonesia. One example is SERVIR that provides mapping for disaster response and has assisted countries in Central America and the Caribbean in responding to hurricanes, earthquakes and other extreme events (GEOSS 2019).

15.4.5 Digital Belt and Road Program—Disaster Efforts

The Digital Belt and Road (DBAR) program and Digital Silk Road Alliance (DSRA) are relatively new activities initiated by the Silk Belt and Road (BAR) initiative. The DBAR is a pioneering international venture to share expertise, knowledge, technologies and data to demonstrate the significance of Earth observation science and technology and applications for large-scale sustainable development projects. The extensive geographical scope of the "BAR" initiative calls for smart uses and applications of big Earth data in the design, development and implementation of diverse projects related to infrastructure improvement, environmental protection, disaster risk reduction, water resource management, urban development, food security, coastal zone management, and the conservation and management of natural and cultural heritage sites. DBAR is committed to implementing projects and actions relevant to the 17 Sustainable Development Goals (SDGs) adopted by the United Nations in September 2015 (United Nations Brussels Team 2018). In the DBAR, natural hazards are an important issue. Belt and Road nations experience approximately 85% of the world's major earthquakes, tsunamis, typhoons, floods, droughts and heatwaves. For example, more than 86,000 people were killed or reported missing in a massive earthquake in Wenchuan, China in May 2008 and the 2004 Indian Ocean earthquake and tsunami killed hundreds of thousands of people. Seven of the top ten countries that saw major losses from disasters between 1995 and 2014 are in this region (Guo 2018, p. 26). The program monitors different types of ecosystems and their evolution, including grasslands, forests, glaciers, urban areas, farmland and coastal regions. Environmental and socioeconomic information will be shared through a platform for big Earth data, scheduled for roll-out between 2016 and 2026. This open-access gateway will allow for researchers, policy makers and the public to track changes, development and trends. The program will investigate indices and indicators to feed into the UN's 2030 Sustainable Development Goals (Guo 2018).

Working group 6 of the DBAR says that DBAR disaster aims to integrate Earth observations (EO) and social vulnerability data to promote implementation of the Sendai Framework in countries along the BAR region. The approach taken by this WG covers satellite information and communication technologies as well as implementation-oriented technologies that involve hardware solutions for risk reduction challenges. "If we do nothing, sensitive environments will be lost and exposure to risks will rise" (Guo 2018).

There are efforts to find solutions using newly defined ideas about big Earth Data. There are four main obstacles to a strategy for the Belt and Road region: poor access to data; a digital divide between developed and developing countries; a lack of awareness of the potential of Earth observations among some policy makers, local scientists and practitioners; and a lack of collaboration. These are long-standing problems—they also slowed emergency responses during and after the Indian Ocean tsunami in 2004, for example.

Important consequences of DBAR strategies necessitate research on new approaches and knowledge improvements. There should be proof of concept for the data. Guo is developing a new concept of big Scientific data and big Earth Data (Guo 2017, p. 4): "Big data is a revolutionary innovation that has allowed the development of many new methods in scientific research. This new way of thinking has encouraged the pursuit of new discoveries. Big data occupies the strategic high ground in the era of knowledge economies and also constitutes a new national and global strategic resource. "Big Earth data", derived from, but not limited to, Earth observation, has macro-level capabilities that enable rapid and accurate monitoring of the Earth, and is becoming a new frontier contributing to the advancement of Earth science and significant scientific discoveries. … Big data research is different from traditional logical research. It uses analytical induction applied to a vast amount of data to statistically search, compare, cluster, and classify. It involves correlation analysis and implies that there may be certain a regularity in the relation between the values of two or more variables; it also aims to uncover hidden correlated networks."

The substantive characteristics of big data computing comprise a paradigm shift from model-driven science to data-driven science, as well as the establishment of a data-intensive scientific approach.

As a branch of big data, scientific big data is a typical representative of data-intensive science. Scientific big data has a number of characteristics, including complexity, comprehensiveness, and global coverage, as well as a high degree of integration with information and communication technology. The approaches used in science are also being transformed—from single-discipline to multidisciplinary and interdisciplinary approaches, from natural science to the integration of natural and social sciences, and from work carried out by individuals or small research groups to projects coordinated by international scientific organizations.

In addition to helping scientists solve hard or previously unsolvable problems through real-time dynamic monitoring and analysis of various related data, the data itself can become an object and tool of research: scientists can conceive, design, and implement research based on the data (Hey et al. 2009 in Guo 2017).

Earth science research, including the atmosphere, land and ocean, has produced huge datasets derived from satellite observations, ground sensor networks, and other sources. This is collectively called big Earth data. Big Earth data has features in common with scientific big data and also has unique characteristics. Big Earth data is characterized as being massive, multisource, heterogeneous, multitemporal, multiscale, high-dimensional, highly complex, nonstationary, and unstructured. It provides support for data-intensive research in the Earth sciences. Modern Earth science requires globally established, quasi real-time, all-weather Earth data acquisition capabilities, and has developed an integrated space-air-ground observation system with high spatial, temporal, and spectral resolutions (Guo 2017).

To realize the above-mentioned efforts, the ISDE organization initiated the Digital Silk Road Alliance (DSRA), established in Sydney in April 2017 with the support of the China Association for Science and Technology (CAST), with the aim of building a network of scientists involved in the Digital Belt and Road initiative and using Digital Earth and geospatial information technologies to solve the scientific problems facing human beings, and to address problems related to the U.N. Sustainable Development Goals.

The DSRA wants to develop Digital Earth in the fields of cartography, remote sensing and geo-information sciences, which are essential for socioeconomic development. Further development of cooperation mechanisms and frameworks toward the development of Earth observation systems and Digital Earth is expected. It is important to use such approaches on global and regional levels in the realms of Earth observation and Digital Earth.

15.4.6 GGIM and DBAR Comparisons and Potential

Comparing the contemporary differences between the U.N. GGIM and the DBAR, the U.N. GGIM is a mature project connected with stable governmental and public infrastructures aiming to address the needs of the SDGs and Sendai DRR and contemporary needs of civil society and its organizations. DBAR has similar ambitions but primarily originated from the countries where spatial data infrastructure (SDI) and national data infrastructure (NSDI) were still not fully developed according to the Silk Belt and Road. The DBAR has a new approach to look for and elaborate big data, mainly based on satellite images. There are still missing concepts regarding delivery of data to interesting groups, the private sector and individual inhabitants (such as the U.N. GGIM using INSPIRE knowledge and experiences). Along the Belt and Road, countries have different political and economic systems and different data, information and knowledge policies. There has been great investment in the DBAR, which created hopes for fast improvement of the situation, but data and information are only part of the efforts, including DRR. In many countries, geoinformatics and cartography are unappreciated. Maps are created without knowledge of how they will be accepted by users (context and adaptive maps) and how the information should be delivered for professionals and public users. This is very important in EW, DRM and DRR.

It is difficult to say which areas will benefit more from Digital Earth. Because the problems are very complex, their solutions require powerful and adaptable tools. Digital Earth is based on integration of various streams and determination of adequate decisions. Informed decisions also rely on the wishes, opinions and reactions of societies, which can be collected via information from social media or volunteers in the field.

It is likely that the main tasks of the U.N. GGIM will be realized incrementally. DBAR activities elaborating important and new aspects of the big data reality will create new situations in data policies in the countries along the Silk Road and Belt. Convergence of both streams will be inevitable and will lead to realization of the dreams of the founders of SDI and NSDI as well as appreciation of modern visualization methods, mainly cartographical ones. Those methods will help experts and the contemporary public to understand problems and cooperate to create solutions for disaster mitigation problems.

15.5 Digital Earth for National and Local Disaster Risk Assessment

Digital Earth is suitable for reporting practices that have been already tested and implemented in one locality and can be successfully adapted in another. Sharing of practices is important in any field of human activity, including disaster risk management. As noted by Amaratunga et al. (2015), sharing of sound practices is intended to improve knowledge sharing for exchange of data and experiences between users on every level—global, national and local.

15.5.1 National Level

The goal of every state is to identify and minimalize risks in its territory. In the Czech Republic, a group of emergency management experts studied the emergency threats and vulnerabilities (Paulus et al. 2016) and identified and categorized the most typical emergency situations. From this analysis, 22 typological emergency situations were pinpointed. A detailed and typified plan for each emergency was defined, including the responsible public administrative organization and the administrative level on which the plans are used (central, regional, local). An indispensable part of each typified plan is the list of recommended spatial data and maps necessary for a successful reckoning of a particular emergency. Public administration bodies are responsible for the development of action plans on the regional and local levels and for identification of key stakeholders.

15.5.2 Local Level—Cities and Urban Areas

A report titled "State of Disaster Risk Reduction at the Local Level: A report on the Patterns of Disaster Risk Reduction Actions at Local Level" (Amaratunga et al. 2015) focuses on disaster risk reduction in urban areas: "Fast growing cities and urban areas of the world increase disaster risk due to economic growth and fast population expansion. … Sound practices that have been tested and implemented by different cities around the world aid knowledge sharing opportunities for future disaster risk reduction. … The intent is to provide local governments and other institutions learn from one another by effectively facilitating the sharing of sound practices and disseminating these established sound practices in risk reduction."

Ten essential goals and examples of well-functioning solutions for local governments to make their cities more disaster-resilient were defined and are listed below (U.N. ISDR 2012).

15.5.3 Existing Methodologies for Risk Assessment

Overviews of how to map and estimate risk have been presented by several scholars (Kappes et al. 2012; Klucka 2014; Forzieri et al. 2016). The European Commission published the Risk Assessment and Mapping Guidelines for Disaster Management (EC 2010), but it was not the first draft of such a pan-European manual. For example, the output of the European project Interreg IIIC Interregional Response to SIPROCI, to which seven countries contributed, is even wider and more thorough than the above mentioned EU final document but was never fully implemented at the European level (SIPROCI 2007). An example of a major non-EU agency that deals with risk discovery and estimation is the Federal Emergency Management Agency (FEMA) from the USA. FEMA announced the release of the State Mitigation Plan Review Guide in 2016 (FEMA 2016) that aids state, tribal, or local governments in developing hazard mitigation plans.

15.5.4 Using Maps for Risk Assessment

An important part of any methodology for identifying and estimating risks is the design of presentations to professionals and the general public. The ideal way to view the risk estimates clearly is a map. The significance and role of maps is described in the book Successful Response Starts with a Map (National Research Council 2007), prepared as a Hurricane Katrina analysis. The creation of maps for risk identification was also described by the above mentioned SIPROCI project (2007) and by other authors including Carpignano et al. (2009) and Winter (1993) described in Dymon (1994). Carpignano et al. (2009) described the development of a decision support system based on a multirisk approach that can overcome difficulties in the overall risk assessment for a territory. To define multirisk maps, a multirisk perspective and stakeholder's perceptions were integrated into a classical risk assessment frame. The specific purpose of this work is to describe the methodological framework built at this stage of the project and discuss the initial results.

Dymon (1994) describes a hazard management map taxonomy offered by Winter (1993) that regards hazard, risk and emergency as the three major categories:

- Hazard maps identify and display the location of hazard zones, areas where there are dangers to humans and their property.
- Risk maps (vulnerability) require calculation of the conditional probability that a given area will experience a particular hazard or a combination of hazards and portray the spatial distribution of those risk computations.
- Emergency maps comprise three additional types: planning, evacuation and crisis maps.

The SIPROCI report (2007) provides a comprehensive method for risk mapping. However, specific proposals were not included in the official final methodology (EC 2010). However, conclusions and recommendations were incorporated into the methodology, such as the by the Fire Rescue Service in the Czech Republic (Krömer et al. 2010), which recommends creating the following types of maps:

- Hazard map—a summary map of the different types of hazards, i.e., a digital map of the manifestations of individual types of emergencies.
- Vulnerability map—the indicator of accumulated vulnerability of the territory as a sum of partial elements of vulnerability.
- Preparedness map—readiness in the territory can be expressed as the availability of forces and means (components of the integrated rescue system) and the availability of means of protection of the population (e.g., coverage of the territory by end elements of the warning).
- Risk map—a summary of all the above map types.

In its official methodology, the EU Risk Assessment and Mapping Guidelines for Disaster Management (EC 2010) only include general recommendations for preparing these types of maps:

- Maps of the spatial distribution of major hazards show the spatial distribution of all relevant elements that need to be protected, such as population, infrastructures, and naturally protected areas.
- The spatial distribution of vulnerability in terms of the susceptibility to damage for all relevant subjects.
- These maps can then provide the basis for the preparation of risk maps in terms of showing the combination of the likelihood and impact of a certain event, as well as for creating of aggregated hazard maps.

However, specific mapping requirements for risk assessment only appear in EU directives for flood mapping such as the Floods Directive (EC 2007). Flood risk mapping is the area of disaster management in which mapping methodologies have advanced the most. The EU directive on the 'Assessment and management of flood risks' requires Member States to conduct an initial assessment for flood hazard maps and flood risk maps:

- The hazard maps should cover geographic areas that could be flooded according to different scenarios. Flood hazard maps show the extent of floods at high- (optional), medium- (at least a 100-year return period) and low-probability floods or extreme events.
- Risk maps should show the potential adverse consequences associated with floods under those scenarios.

15.5.5 The Benefits of Digital Earth for Risk Assessment—Using Dynamic Data

Creating maps with the standardized content and symbolism mentioned in the previous section is necessary for preparing the components of an integrated rescue system for crisis situations and for managing them. However, the basis of the Digital Earth concept is not the creation of printed and static maps, but the dynamic sharing of different types of data, including near-real time data sharing. The development of electronics, networks, databases, data sharing (included in Digital Earth) brings new possibilities for risk assessment.

As an example, for a risk assessment at a particular location and at a certain time, it is possible to take advantage of the current location of mobile phones, from which the present population can be estimated more accurately than using the standard census data. Extensive studies focused on different aspects of human presence estimation based on mobile phone data have been presented, particularly from Europe and Asia (Ahas et al. 2010; Batista e Silva et al. 2013; Cao et al. 2017; Järv et al. 2017; Kang et al. 2012). Kubíček et al. (2018) proposed analysis of human presence using data from mobile operators. The analysis is based on a dataset describing the estimated human presence (EHP) with two values—visitors and transiting persons—depending on the overall time spent within a specific mobile cell.

The advantage of using the EHP numbers over data from a census was analyzed during the Integrated Rescue System training held in 2017 in Brno, Czech Republic. The goal was to decide where to locate water tanks with supplies of drinking water for inhabitants in case the standard water supply network becomes contaminated.

This emergency situation is demonstrated in Fig. 15.3. "The location-allocation analysis on the leftmost side only takes into account census data and evenly distributed population throughout the administrative unit. Each water tank can supply approximately 2000 people. The second analysis adjusts the water tank locations according to the real locations of buildings and population in administrative units. The third and fourth analysis quantifies the EHP for working days and weekends. Using of EHP proposes a greater number of necessary water tanks in administrative units, and their optimal locations change according population fluctuations" (Kubíček et al. 2018).

Risk assessment is addressed at different levels (international, national and local), and each of these levels has its own goals and uses. It is very useful to share experiences and data between these levels. This allows for generalization of knowledge and results from the local level to the national and international levels. Such analyses can become an engine for developing better risk assessment methods and disaster mitigation. The Digital Earth concept linking databases and enabling data sharing provides a methodological and technological background for this goal.

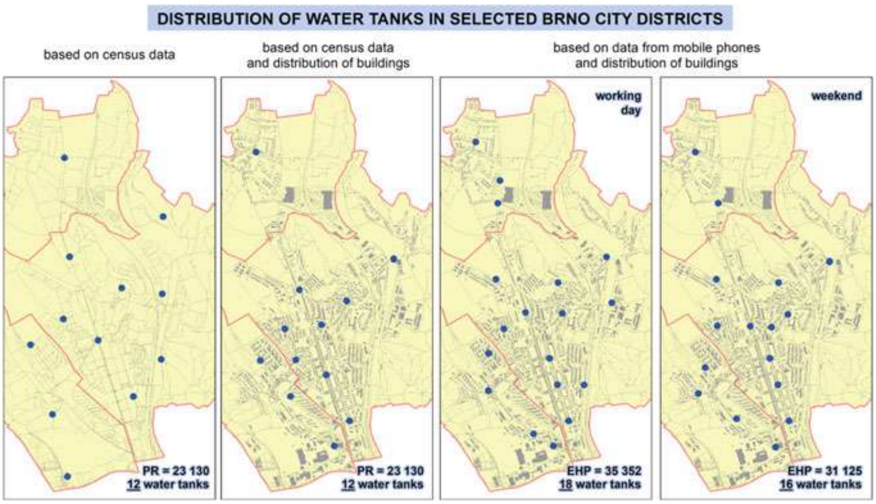


Fig. 15.3 The role of the spatiotemporal distribution of the population in the case of a water shortage. Reprinted from Kubíček et al. (2018) by permission of Taylor & Francis Ltd.

15.6 Digital Earth and Disaster Risk Mapping Technology

15.6.1 The Role of Cartography in Disaster Risk Mapping

In the frame of disaster risk mapping, geographic knowledge is crucial for making proper decisions. The importance of spatial information and its potential support for emergency actions were stressed and evaluated by several authors (Kevany 2008; Zlatanova and Li 2008; and Konecny 2006). Among the various ways to transmit, share, and visualize geographic knowledge, cartography is one of the most important. Cartography and geoinformatics have experienced a huge technological shift over the last 30 years. Digital Earth systems have become important foundations for data management related to geographic phenomena.

The application of dynamic cartographic visualization opens the possibilities of adaptive cartography. It allows for creating maps of current risks (e.g., current and predicted flooded area or direction of fire spread), the location of nearby emergency services, or escape routes for the population at risk.

The theory of using adaptive cartography for emergency management geographic support was described by Reichenbacher (2003) and Meng (2005). This method is based on the idea of geographic data visualization automation and adjustment according to the situation, purpose and user's background (Reichenbacher 2003).

The adaptation of maps can generally be defined by a number of "Ws"—what, when, where, who, and how—as documented in Fig. 15.4. It illustrates the types of contexts that can influence the conditions of disaster risk mapping.

		type of context	
What ?	What happend?	*SITUATION*	**EMERGENCY CONTEXT**
	What needs to be done?	*ACTIVITY*	
When?	When the event occured?	*TIME*	
	When the activity is realised?	*PHASE*	
Where?	Where the event occured?	*LOCATION*	
	(What area) is affected by the event? (What) is the extent of the activity?	*OPERATIONAL RANGE*	
Who?	Who is the user of the map?	*USER ABILITY*	
	Who is the data manager?	*DATA MANAGEMENT*	
How?	How the map is used?	*MAP FUNCTION*	
	(What) is the size of the display?	*TECHNOLOGY*	

Fig. 15.4 Possible contexts influencing map use and mapping. Adapted from Kozel et al. (2011)

15.6.2 Use Case Examples

The adaptive mapping principles described in the previous section were demonstrated in several scenarios, e.g., Talhofer et al. 2007, Mulickova et al. 2007. One of the scenarios, called "FLOOD", aims to improve flood management. A case study was practically verified in the winter of 2011/2012, when one of the field experiments was performed. Based on an analysis of the flood management system in the Czech Republic (Kubíček et al. 2011), five main ACTIVITIES were defined for the Flood Use case (SITUATION):

- PREDICTION AND PROGRESS—development and expected progress of the flood
- TECHNICAL SUPPORT—technical support in the inundated area—support of Flood Security Activities
- RESCUE—the evacuation of citizens
- ORGANIZATION—an organization of powers and means
- PUBLIC INFORMATION—information for the public on flood development, evacuation, etc.

Some of the ACTIVITIES defined above are universal (e.g., organization) and may be performed in different SITUATIONS whereas others (e.g., flood prediction) are situation-specific.

There were a few principal operational ranges defined in the presented use case: FLOODPLAIN for detailed information on the inundation, REGION/DISTRICT/MUNICIPALITY to comply with the hierarchical order of the flood management system, CATCHMENT to monitor the flood at natural borderlines, and SECTION for a detailed view of the municipality.

15.6.3 Use Case Adaptation Principles

The fact that an object is evaluated from the perspective of a defined context is fundamental to the map symbol adaptation process. The most important aspect of the geographic feature may not be the character of the object as defined by the data source, but what ROLE it plays in the decision-making process. The map symbol is an expression of such a ROLE. Because the data are typically collected for purposes other than emergency management, semantic relations must be defined, and new roles should be specified.

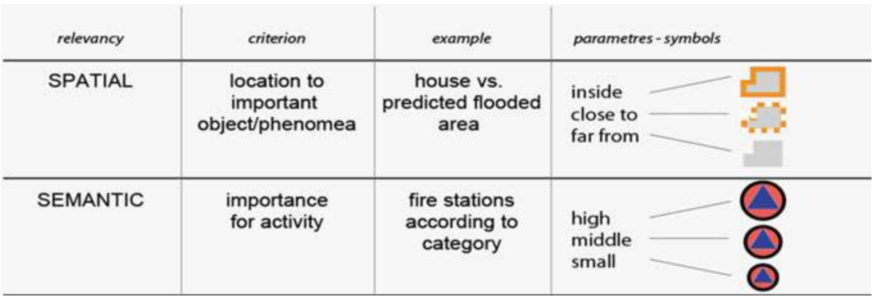
Based on context, the semantic relevance is assessed. Information on the geographic object is relevant if it is necessary for the decision-making process within the context. The relevancy assessment is important from the cartographic point of view since the large number of objects that are visualized on the map limit its legibility and thus the effectiveness of the cartographic visualization as a decision support tool.

When information is relevant, we can assess the degree of relevancy and use other cartography means to increase/decrease the importance of a spatial object or phenomena. The relevancy degree can be assessed for both the semantic and spatial aspects, as illustrated in Fig. 15.5.

The activity and the crisis event itself undergo temporal changes and thus the object properties change as well. For example, if the water level is rising and another house is endangered or a house is already evacuated. These facts should be considered during map symbol design.

15.6.4 Context Map Composition

The process of data model definition is illustrated in Fig. 15.6. The emergency context defines the basic data model (e.g. the information content of the map), and relevant



relevancy	*criterion*	*example*	*parametres - symbols*
SPATIAL	location to important object/phenomea	house vs. predicted flooded area	inside close to far from
SEMANTIC	importance for activity	fire stations according to category	high middle small

Fig. 15.5 Degrees of spatial and semantic relevance. Cartographic symbols prepared by L. Friedmannova. Adapted from: Brezinova et al. (2011)

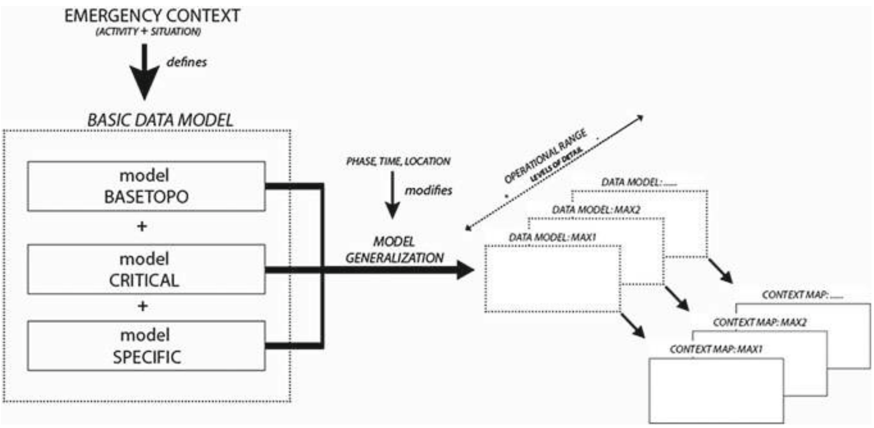


Fig. 15.6 Data model definition. Adapted from Mulickova (2011)

features of the models BASETOPO, CRITICAL, and CONTEXT SPECIFIC are selected. The basic model is then modified as the context is more precisely specified (i.e., according to the PHASE). The model is generalized and further specified for each level of detail within the operation range.

The examples of a context map for flood management in Fig. 15.7 document different context views of the spatial database. Context maps for three emergency contexts—PREDICTION (A), RESCUE (B, D) and ORGANIZATION (C) are shown. The level of detail corresponds with the operation range "section". Maps share the same topographic background (i.e., BASETOPO) and, to a certain extent, flood-SPECIFIC features (i.e., the flood extent and buildings in it). The visualizations differ in activity-specific features—i.e., features specific to prediction (flood activity degree, number of affected persons), to the organization (places of intervention and its description) and to the rescue efforts (evacuation zones, routes). The features of the CRITICAL model are not included.

Maps A and B in Fig. 15.7 illustrate the phase of preparation—there is no flooding yet but there is a prediction of flooding. At that time, houses are endangered. In the response phase (Maps C and D), houses are already affected. The visualization changes are based on the progress of the disaster event.

Maps B and D support the same activity (i.e., rescue) but in different phases. The maps display visualization changes based on the progress of the activity. In the preparation phase, the zone of evacuation is marked and buildings for evacuation are selected. In the response phase, all the buildings are already evacuated.

A possible technical implementation is described in detail by Kozel (2009) and Kozel and Štampach (2010).

15.6.5 Existing Symbol Systems for Disaster Management

Cartography plays a key role in disaster management for a clear representation of the necessary objects and phenomena to decision makers. Upon the occurrence of disasters, crisis management actors need specialized maps to provide a clear idea of the emergency, localization, distribution and characteristics. One of the objectives of cartographers is to design effective representation of spatial information through graphic symbols (Akella 2009). The symbols should indicate information about depicted objects and phenomena without the use of a legend, especially in an emergency. They should also provide users qualitative and quantitative information for the presented object or phenomenon (Konecny and Bandrova 2006).

A number of agencies and organizations related to disaster protection have developed databases, geo-portals and cartographic products for crisis management and adopted their own standards for symbols.

One of the most popular symbol systems for crisis management is the set of 500 humanitarian symbols of the United Nations Office for the Coordination of Humanitarian Affairs (OCHA). The symbols are freely available at http://reliefweb.int/ and aim to help disaster responders present information about crisis situations

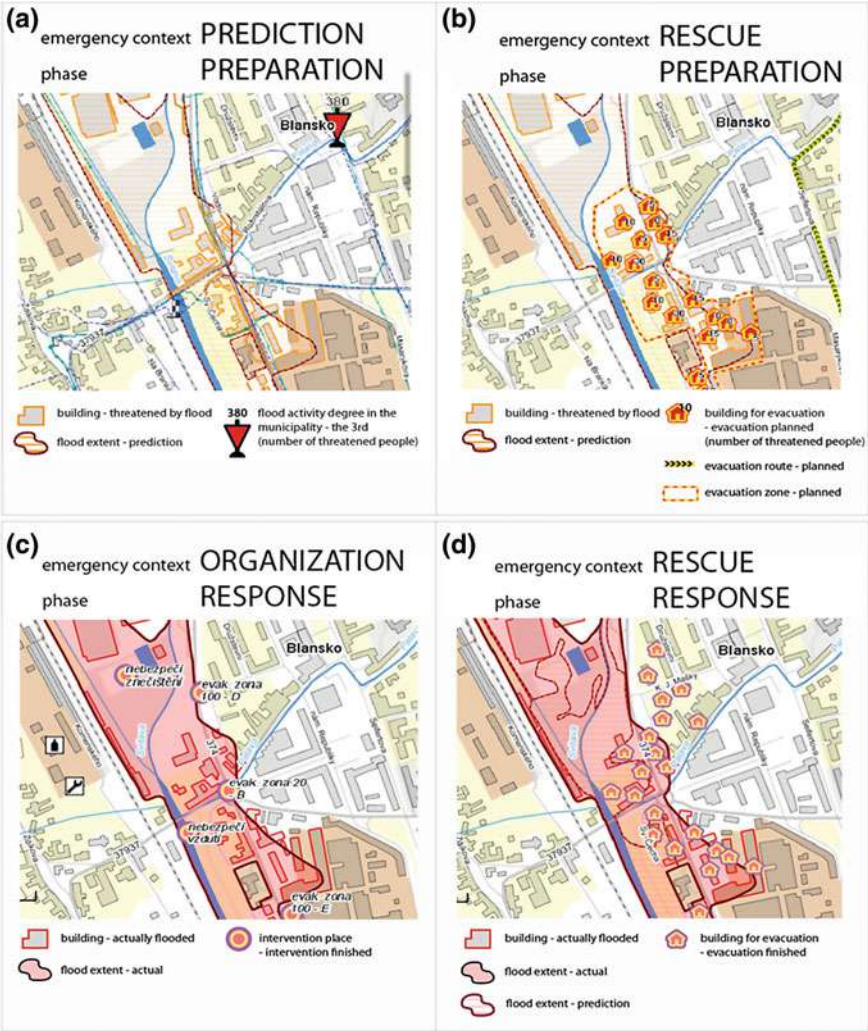


Fig. 15.7 Examples of context mapping for various phases of the emergency management cycle. *Source* Mulickova and Kubicek (2011)

quickly and simply (United Nations Office for the Coordination of Humanitarian Affairs 2012). The symbols can be used to produce humanitarian reports, maps, and websites. The OCHA humanitarian icons are divided into 17 categories. The set of symbols covers both disasters and activities, including the supply of water containers and equipment shelter, access to people in need and protection of civilians. The icons are associative and have a simple structure that allows for easy comprehension.

The Emergency Mapping Symbology (EMS) in Canada was developed under the auspices of GeoConnections, with participation from emergency management organizations across Canada. It was designed to be used by federal, provincial, regional and local organizations involved in the management of major events, disasters, and other incidents where emergency help and security are needed (GeoConnections 2010). The EMS contains a set of symbols and a four-level, hierarchical classification of the entities. The categories include incidents, infrastructures, operations, and aggregates. Symbols in the same category have similar colors. There is also a second version of the symbols adapted for black and white printing.

The Association of Volunteer Emergency Response Teams developed a project called Disaster Response Map Symbols (DRMS) as an effort to compile a standard set of symbols aimed to support the creation of efficient maps for disaster management. It comprises 285 symbols. The DRMS contains 5 families of symbols in a single font, including vehicles, infrastructure, mobile/temporary services and teams, events, ships and some special symbols (Association of Volunteer Emergency Response Teams 2009).

Another popular symbol system is the symbology developed by federal, state, and local agencies in the USA working together under the auspices of the Federal Geographic Data Committee (FGDC) Homeland Security Working Group. The symbol system includes symbols and their definitions for the categories of incidents, natural events, operations, and infrastructures. The structure of each category and a damage-operational status hierarchy were developed using color and frame shapes with line patterns (Homeland Security Working Group 2017). The symbols are designed to be presented in color or black and white formats.

The cartographic symbols should have clear and short definitions to be used in a map legend. One very important characteristic is that they are situated on a map and should indicate qualitative and quantitative information about the represented object, phenomena or process to users.

Considering the advantages and disadvantages of existing emergency symbol systems, a new symbol system for the needs of disaster management was developed at the Laboratory on Cartography of the University of Architecture, Civil Engineering and Geodesy in Sofia. The Symbol System for Disaster Management (SSDM) was developed to support thematic mapping for early warning and crisis management and operational activities of all participants in disaster management, as well as to help citizens understand specialized emergency maps. The SSDM was designed to be useful for the general public as well as for professionals.

15.6.6 Opportunities for New Disaster Risk Mapping Technologies

The technological shifts in cartography and geoinformatics were on the level of data analysis and visualization, bringing new data sources from different sensors and mapping strategies. One of the most notable examples of this is cell phone data.

Data derived from active cell phones or active SIM cards for some administrative units are becoming available for various uses (see an example from the Czech Republic, the O2 Liberty API, https://www.o2.cz/podnikatel/liberty-api/). Analysis of the number of SIM cards and existing demographic data has opened a novel set of possible applications for emergency management and disaster risk mapping. The availability of cell phone data enables the following:

- More accurate estimation of the actual number of inhabitants within the administrative unit and their temporal rhythms (example on Brno, Czech Republic in Kubíček et al. 2018). Comparing such an analysis with the existing census data and annual demographic reports (see Fig. 15.8), the administrative units can be further divided into several typological units (with the maximum during working days, weekends, etc.) In addition, the population estimations can be used to better plan the evacuation and other inhabitant-sensitive activities during emergencies.
- The cell phone data analysis often reveals regular trends as described above and some irregular peaks and peculiarities. These high concentrations of inhabitants are connected with cultural and sports events such as concerts and music festivals.

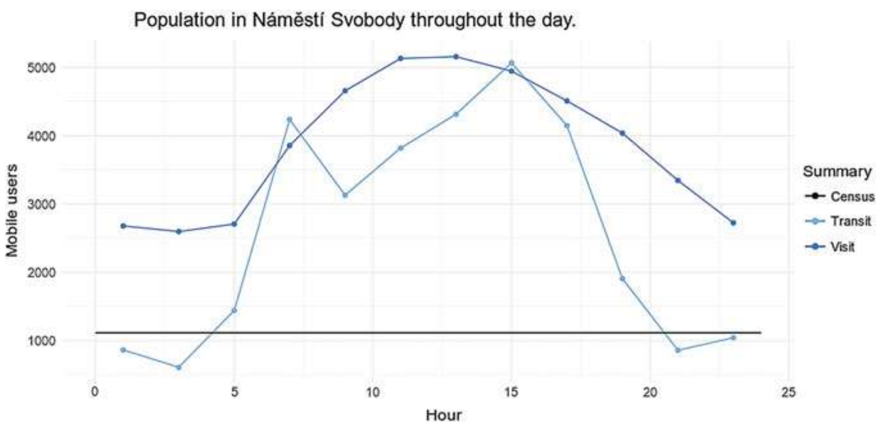


Fig. 15.8 Variability of the population in an administrative unit Náměstí Svobody, Brno, Czech Republic. Comparison of cell phone and census data. Reprinted from Kubíček et al. (2018) by permission of Taylor & Francis Ltd.

15.6.7 *Future Directions—New Symbol System for Disaster Management (SSDM)*

The examples, approaches and case studies described above provide various opportunities for future development and applications such as the development of virtual and augmented reality tools and devices. The Digital Earth concept can be also understood as a virtual reality system (Çöltekin et al. 2019).

The new cartographic Symbol System for Disaster Management (SSDM) was created in Bulgaria after proposing a classification structure of represented objects and phenomena, construction and design of symbols, implementation in real situations and use in map compiling for disaster preparedness.

15.6.7.1 Classification Structure

The SSDM consists of a 4-level hierarchical classification of objects and phenomena concerning disaster management and a set of 115 symbols. At the highest level, the objects are divided into 5 categories: disasters, infrastructure, protection services and safety infrastructure, affected people and infrastructure, and operational sites and activities. Each category is divided into classes, which are divided into subclasses that consist of objects and phenomena (Fig. 15.9).

15.6.7.2 Design of Symbols

The ability of symbols to transmit information and the way they are perceived by map users are critically important. The design process of the SSDM started with consideration of the rules of construction and use of symbol systems, examination of the relations between objects and phenomena, their classification and specifics.

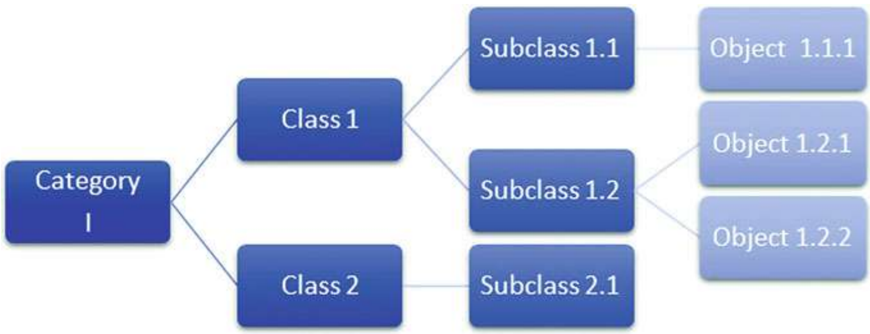


Fig. 15.9 Classification structure *Source* Marinova (2018)

The design was accompanied by optimal requirements to achieve readability, expressiveness and visibility, taking into account modern technologies and techniques in cartography. It is challenging to choose graphical variables so that all the symbols can be quickly and easily perceived and are associative and properly referred to their respective categories.

All categories of the SSDM are distinguishable by their shape and color. The symbols consist of white pictograms and shapes with various background colors. The choice of background colors, except to achieve clear distinctiveness, depends on the message that the symbols should express to the users. A psychological perception of the colors was taken into account. The different shapes for the categories aim to avoid potential problems resulting from low light or black and white printing.

Each category has an individual letter code for easy identification: A—disasters; B—infrastructure; C—protection services and safety infrastructure; D—affected people and infrastructure; and E—operational sites and activities. Each object and its respective symbol have an alphanumeric code formed by the category code and the serial number of the object in its category.

Figure 15.10 presents part of the symbol system, including the alphanumeric code, graphic symbol and a brief description.

The status of objects in "infrastructure" and "protection services and safety infrastructure" in a crisis situation is represented by a combination of symbols in category B (infrastructure) and category C (protection services and safety infrastructure), with symbols representing destroyed, affected and unaffected objects of category D shown in a reduced size (Fig. 15.11).

15.6.7.3 Maps for Disaster Protection

The new Symbol System for Disaster Management was applied in experimental development of training maps supporting actions in emergencies and in a series of maps for disaster protection at local and regional levels. The main tasks of local and regional disaster protection plans are the analysis and assessment of disaster risks, prevention and mitigation, early warning, and coordination of disaster management activities. Participants in these activities need specialized geographic information to support concrete actions.

The SSDM was applied in the production of base maps of the municipality of Troyan, Bulgaria, at a scale of 1:50000 (Fig. 15.12) and Troyan at a scale of 1:10000 (Fig. 15.13). The maps were compiled according to predefined elements of map content and aim to support activities described in the disaster protection plan of the municipality.

The main features of hydrography, settlements, infrastructure (including transport, telecommunication, energy, manufacturing and water infrastructure) as well as services and facilities related to disaster protection (such as hospitals, shelters, and helicopter pads) are represented by the SSDM. Based on the main disaster protection maps, a series of maps for disaster management in case of earthquakes, floods, fire, and industrial accidents were created. Additional information was provided for

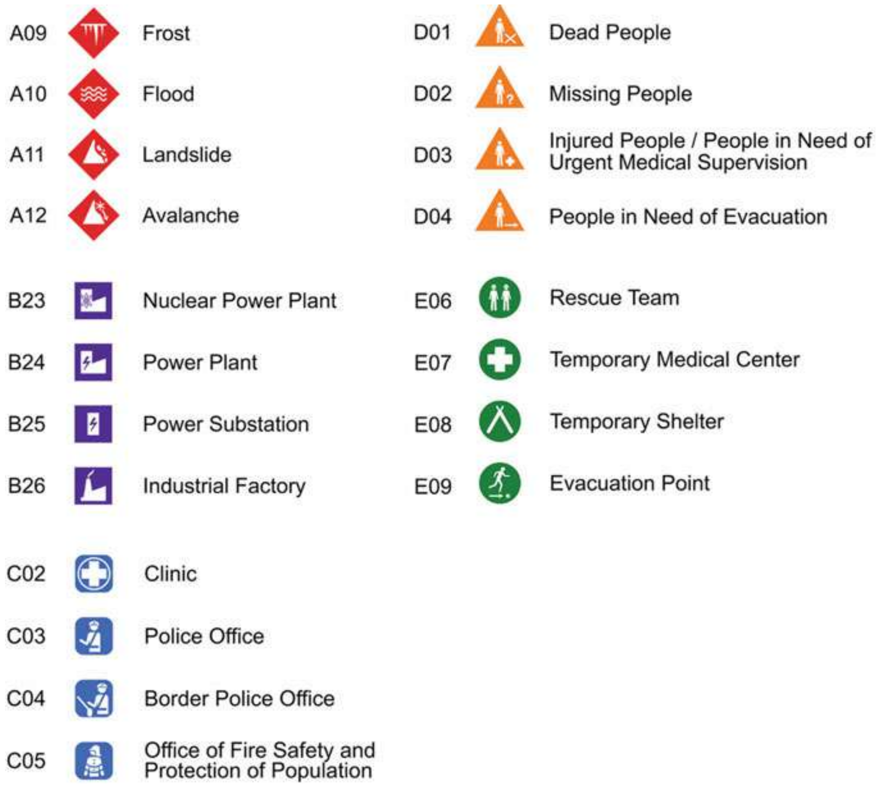
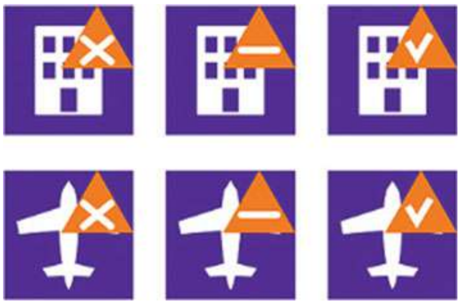


Fig. 15.10 Symbols of the symbol system for disaster management (SSDM) *Source* Marinova (2018)

Fig. 15.11 Symbols to present affected infrastructure *Source* Marinova (2018)



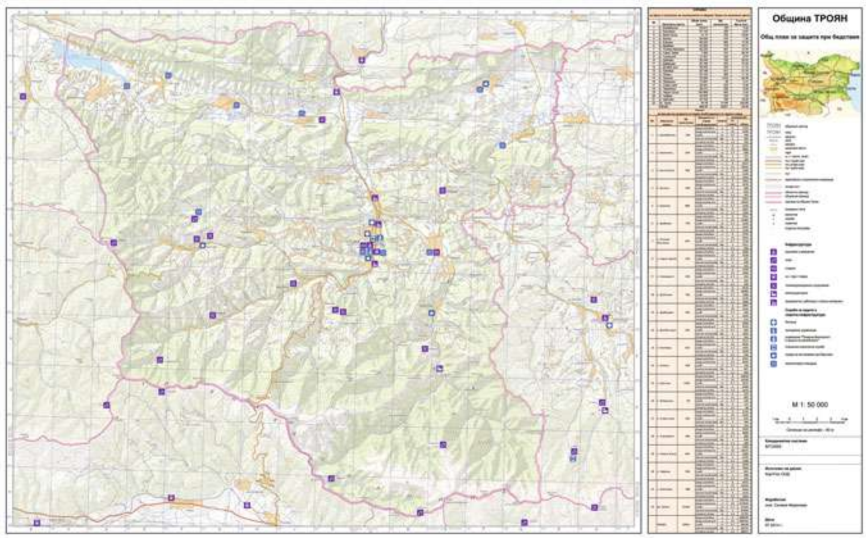


Fig. 15.12 Base map for disaster protection *Source* Marinova (2018)

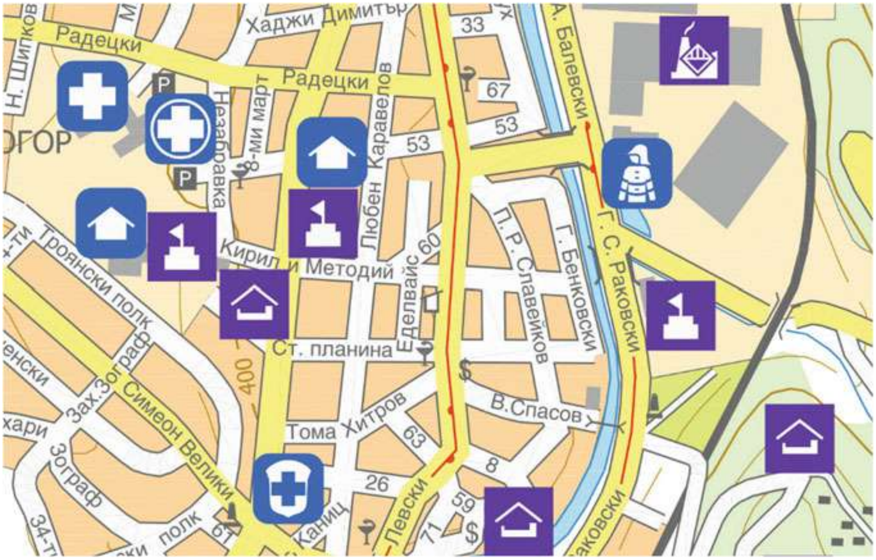


Fig. 15.13 Base map for disaster protection (partial). *Source* Marinova (2018)

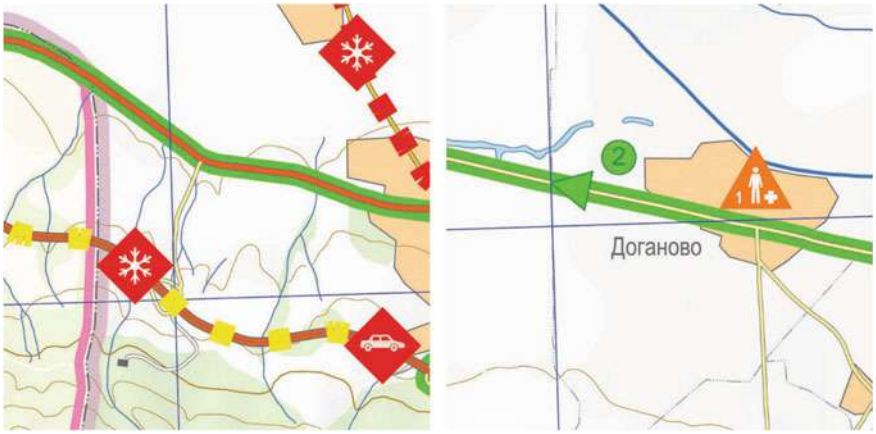


Fig. 15.14 A map for evacuation planning Adapted from: Marinova (2018)

some features including the object name and description, number of beds in shelters, dangerous industrial objects, type of stored materials, and fire-fighting equipment. Infrastructure and services/facilities for protection are represented by the symbols in Category B and Category C (Fig. 15.13). These maps also support predisaster activities, including assessment and preparedness.

In a crisis situation base maps can be processed into rapid and reference maps presenting the type and location of disaster(s) by adding symbols from Category A and symbols for affected people and affected infrastructure in Category D (Fig. 15.14). The symbols for operational sites and activities (Category E) could be useful for damage assessment and recovery in the postdisaster stage.

The map content and displayed information of operational situations could help support the responsible authorities and individuals to make timely and effective decisions. Such maps could allow for identification of the affected areas in municipalities or regions, and provide significant contributions to population protection, mitigation and evacuation planning operations.

Cartography plays a key role in the main stages of disaster management. Efficient and cooperative preventive and protective activities of authorities require appropriate and easily understood geographic information. The use of a standard system of associative symbols can facilitate significantly cooperative disaster management strategies at local, regional and international levels.

15.7 Conclusion

Disaster mitigation and DRR are complicated processes, and solutions could be improved by using powerful tools such as Digital Earth. The concept of DE covers almost all activities occurring in ICT in the contemporary world. To be successful in employing the right solutions, we need to create improved concepts that consider the newest knowledge about disaster mitigation and DRR. To realize this, we need well-organized data and information such as in data ecosystems (as in the U.N. GGIM) that reflect the complexity of the problems to be solved, defined by the SDGs. Sharing data and information, visualization with the help of digital maps, cartographic models and their combinations hold important promise to support decision makers and society with true and understandable outputs to help to comprehend situations, to create instructions and standards on how to behave in various situations, and to be ready when risks transform into disasters. This chapter highlighted the newest projects, including the U.N. GGIM and DBAR. In the future, these approaches with commonalities and differences should be developed to support smart solutions for human society.

References

Ahas R, Silm S, Järv O et al (2010) Using mobile positioning data to model locations meaningful to users of mobile phones. J Urb Technol 17(1):3–27

Akella MK (2009) First responders and crisis map symbols: clarifying communication. Cartogr Geogr Inf Sci 36(1):19–28

Amaratunga D, Pathirage C, Malalgoda C (2015) State of disaster risk reduction at the local level: a report on the patterns of disaster risk reduction actions at local level. Research report. UNISDR, Geneva, Switzerland

Association of Volunteer Emergency Response Teams (2009) Disaster response map symbols, version 1.01 (Beta). Association of Volunteer Emergency Response Teams, Horsham, UK

Brezinova S, Drapela MV, Friedmannova L et al (2011) Producing of cartographic infrastructure. In: Konecny M (ed) Dynamicka geovizualizace v krizovem rizeni. Masaryk University, Brno, Czechia, pp 235–287

Cao J, Tu W, Li Q et al (2017) Spatio-temporal analysis of aggregated human activities based on massive mobile phone tracking data. J Geoinf Sci 19(4):467–474

Carpignano A, Golia E, Di Mauro C et al (2009) A methodological approach for the definition of multi-risk maps at regional level: first application. J Risk Res 12(3–4):513–534

Çöltekin A, Oprean D, Wallgrün JO et al (2019) Where are we now? Re-visiting the Digital Earth through human-centered virtual and augmented reality geovisualization environments. Int J Digit Earth 12(2):119–122

De Longueville B, Annoni A, Schade S et al (2010a) Digital Earth's Nervous System for crisis events: real-time Sensor Web Enablement of Volunteered Geographic Information. Int J Digit Earth 3(3):242–259

De Longueville B, Annoni A, Schade S et al (2010b) Digital earth's nervous system and volunteered geographic information sensing: towards a self-aware planet. In: Abstracts of 3rd ISDE digital earth summit. ISDE, Nessebar, Bulgaria, 12–14 Jun 2010

Dymon UJ (1994) Mapping-The Missing Link in Reducing Risk under SARA III, RISK: Health, Safety & Environment 5(4):337–349

e Silva FB, Gallego J, Lavalle C (2013) A high-resolution population grid map for Europe. J Maps 9(1):16–28

EC (2007) Directive 2007/60/EC of the European parliament and of the council of 23 October 2007 on the assessment and management of flood risks

EC (2010) Risk assessment and mapping guidelines for disaster management. Commission staff working paper

EC (2019) Emergency management service - service overview. https://emergency.copernicus.eu/mapping/sites/default/files/files/CopernicusEMS-Service_Overview_Brochure.pdf. Accessed 21 Jul 2019
FEMA (2016) State mitigation plan review guide, FP 302-094-2. https://www.fema.gov/media-library/assets/documents/101659. Accessed 13 May 2019
Forzieri G, Feyen L, Russo S et al (2016) Multi-hazard assessment in Europe under climate change. Clim Change 137(1):105–119
GeoConnections (2010) Emergency mapping symbology, version 1.0. http://gisuser.com/2010/06/geoconnections-announces-emergency-mapping-symbology-now-available/. Accessed 15 Mar 2014
GEOSS (2019) GEOSS. www.earthobservations.org/geoss. Accessed 2 Jun 2019
Goodchild MF (2008) The use cases of digital earth. Int J Digit Earth 1(1):31–42
Gore A (1998) The digital earth: understanding our planet in the 21st Century. Aust Surv 43(2):89–91
Guo H (2017) Big earth data: a new frontier in Earth and information sciences. Big Earth Data 1(1–2):4–20
Guo H (2018) Steps to the digital silk road. Nature 554:25–27
Hey T, Tansley S, Tolle K (2009) The forth paradigm: data intensive scientific discovery. Microsoft Research, Washington, DC
Homeland Security Working Group (2017) Symbology reference. http://www.fgdc.gov/HSWG/index.html. Accessed 20 Feb 2017
Järv O, Tenkanen H, Toivonen T (2017) Enhancing spatial accuracy of mobile phone data using multi-temporal dasymetric interpolation. Int J Geogr Inf Sci 31(8):1630–1651
Kang C, Liu Y, Ma X et al (2012) Towards estimating urban population distributions from mobile call data. J Urb Technol 19(4):3–21
Kappes MS, Keiler M, von Elverfeldt K et al (2012) Challenges of analyzing multi-hazard risk: a review. Nat Hazards 64(2):1925–1958
Kevany MJ (2008) Improving geospatial information in disaster management through action on lessons learned from major events. In: Zlatanova S, Li J (eds) Geospatial information technology for emergency response. Taylor & Francis Group, London, UK, pp 3–19
Klucka L (2014) Mapping of security risk in the South Moravia District. MSc Thesis. Masaryk University, Faculty of Science, Department of Geography, Brno, Czechia
Konecny M (2006) Mobile and adaptive cartography and geoinformatics in early warning and crises management. In: Seventeenth United Nations regional cartographic conference for Asia and the Pacific. UNO, Bangkok, p 12
Konecny M, Bandrova T (2006) Proposal for a standard in cartographic visualization of natural risks and disasters. Int J Urb Sci 10(2):130–139
Konecny M, Zlatanova S, Bandrova T (2010) Geographic information and cartography for risk and crisis management. Towards better solutions. Springer, Berlin Heidelberg
Kozel J (2009) Contextual map service. PhD Dissertation. Masaryk University, Faculty of Science, Department of Geography, Brno, Czechia
Kozel J, Ludik T, Mulickova E et al (2011) Principy dynamicke geovizualizace. In: Konecny M (ed) Dynamicka geovizualizace v krizovem rizeni. Masaryk University, Brno, Czechia, pp 99–128
Kozel J, Štampach R (2010) Practical experience with a contextual map service. In: Konecny M, Zlatanova S, Bandrova TL (eds) Geographic information and cartography for risk and crisis management: towards better solutions. Springer, Berlin, Heidelberg, pp 305–316
Krömer A, Musial P, Folwarczny L (2010) Mapovani rizik. Spektrum, Ostrava, Czechia, 126 p.
Kubíček P, Konečný M, Stachoň Z et al (2018) Population distribution modelling at fine spatio-temporal scale based on mobile phone data. Int J Digit Earth 1–22 https://doi.org/10.1080/17538947.2018.1548654
Kubíček P, Muličková E, Konečný M et al (2011) Flood management and geoinformation support within the emergency cycle (EU Example). In: Hřebíček J, Schimak G, Denzer R (eds) Environmental software systems. Frameworks of eEnvironment. Springer, Berlin, Heidelberg, pp 77–86

Marinova S (2018) Thematic mapping and visualization for early warning and crisis management. UACEG, Sofia. ISBN 978-954-724-108-4
Meng L (2005) Egocentric design of map-based mobile services. Cartogr J 42(1):5–13
Mulickova E (2011) Cartographic models. In: Konecny M (ed) Dynamicka geovizualizace v krizovem rizeni. Masaryk University, Brno, Czechia, pp 210–234
Mulickova E, Kozel J, Kubicek P (2007) Utilisation of contextual visualization in monitoring of dangerous goods transportation. In: Interop-soft 2007. MSD Brno, Brno, pp 47–53
Mulickova E, Kubicek P (2011) Adaptive mapping in flood management. In: Konecny M, Mulickova E, Kubicek P et al (eds) Geoinformation support for flood management in china and the Czech Republic. Masaryk University, Brno, Czechia, pp 91–101
National Research Council (2007) Successful response starts with a map: improving geospatial support for disaster management. The National Academies Press, Washington, DC
Paulus F, Krömer A, Petr J et al (2016) Threats analysis for the Czech Republic. IOOL, Lazne Bohdanec
Reichenbacher T (2003) Adaptive methods for mobile cartography. In: 21st International cartographic conference, ICA, Durban, South Africa, 10–16 August 2003
Scott G (2018) UN-GGIM: strengthening the global data ecosystem. https://un-ggim-europe.org/wp-content/uploads/2018/11/4-UNGGIM-Presentation-GScott-6June2017.pdf. Accessed 10 May 2019
SIPROCI (2007) Interregional response to natural and man-made catastrophes. http://www.cti.gr/en/activities-en/development-projects-en/item/164-siproci/164-siproci. Accessed 14 May 2019
Talhofer V, Kubicek P, Brazdilova J et al (2007) Dynamic cartographic visualisation in a process of transportation monnitoring of dangerous chemical substances. In: Proceedings of the International conference on military technologies 2007. University of Defence, Brno
U.N. ISDR (2009) Terminology. https://www.unisdr.org/we/inform/terminology. Accessed 10 May 2019
U.N. ISDR (2012) How to make cities more resilient: a handbook for local government leaders. United Nations, Geneva
United Nations (2016) Conference on housing and sustainable urban development (Habitat III), The new urban agenda, Quito, Ecuador, 2016. http://habitat3.org/wp-content/uploads/NUA-English.pdf. Accessed 20 Apr 2019
United Nations Brussels Team (2018) The sustainable development goals (SDGs). https://www.unbrussels.org/the-sustainable-development-goals-sdgs. Accessed 15 May 2019
United Nations General Assembly (2015) Sendai framework for disaster risk reduction 2015–2030. In: The third World conference on disaster risk reduction, United Nations, Sendai, Japan, 3 June 2015
United Nations Office for the Coordination of Humanitarian Affairs (OCHA) (2012) World: humanitarian and country icons 2012. http://reliefweb.int/report/world/world-humanitarian-and-country-icons-2012. Accessed 23 Feb 2017
WCDRR (2016) World conference for disaster risk reduction. https://en.wikipedia.org/wiki/World_Conference_on_Disaster_Risk_Reduction. Accessed 10 May 2019
Winter NL (1993) Hazard management mapping: a taxonomy. In: 16th Proceedings of appllied geography conference
Zlatanova S, Li J (2008) Geospatial information technology for emergency response. Taylor & Francis Group, London, UK

Milan Konecny is a Professor at Masaryk University, Czech Republic. He was/is a Guest Professor of the Chinese University of Hong Kong, the Henan University, the Shenzhen University, the Vienna University, etc.; a President of International Cartographic Association; Acting President and the Vice President of International Society for Digital Earth and Academician and Vice-President of International Eurasian Academy of Science.

Temenoujka Bandrova is President of Bulgarian Cartographic Association and head of Laboratory on Cartography, University of Architecture, Civil Engineering and Geodesy, Sofia. She is a Vice-President of ICA and treasurer, bureau member of ISDE. She is an author of 40 school atlases and more than 100 wall maps.

Petr Kubicek is an Associated Professor at the Masaryk University. He specializes in digital cartography and geoinformation infrastructure. Kubicek cooperates with the International Cartographic Association commissions on Use, User and Usability Issues and Cognitive Issues in Geographic Information Visualization. He is a member of the national GeoInfo Strategy team.

Silvia Marinova is Secretary General of Bulgarian Cartographic Association and Chief Assist. Prof. at University of Architecture, Civil Engineering and Geodesy, Sofia. She also serves as a cartographer at KartGeo. Her science interests are focused on thematic mapping for disaster management, cross-border mapping and mountain cartography.

Radim Stampach Ph.D. is a lecturer and GIS specialist at the Masaryk University. He is interested in using VGI data and maps in crisis management and for humanitarian purposes.

Zdeněk Stachoň Ph.D. works as an assistant professor on the Department of Geography, Masaryk University. He focuses mainly on the principles of map design, cognitive aspects of cartographic visualization, and the geographic support of decision making. He is a member of the Czech Cartographic and Geographic societies, member of editorial board of M.A.P.S. edition.

Tomáš Řezník is an Associated Professor of Cartography, Geoinformatics and Remote Sensing at the Department of Geography and a vice-head at the Laboratory on Geoinformatics and Cartography (LGC) at Masaryk University. He worked as a post-doctorate scientific/technical project officer at the Joint Research Centre (JRC) of the European Commission, within Digital Earth and Reference Data unit (currently Digital Economy unit).

Chapter 16
Digital City: An Urban Perspective on Digital Earth

Davina Jackson and Richard Simpson

Abstract Digital Earth and many other satellite and semiconductor-enabled cartography advances imply the need for a globally useful schema for more scientific and eco-ethical management of cities. How should we plan an internationally cohesive and locally effective system for understanding and managing urban stocks and flows around our planet? The answer to this question depends on new systems for managing geodata to underpin increasingly automated systems for evidence-based decision making. The current concept of Digital Earth as a "self-aware nervous system" is being advanced by urban proto-projects that are supported or followed by globally applicable initiatives including Singapore's new Geospatial Masterplan, the International Standards Organization's City Standards, Denmark's Open Public Life Data Protocol, and the City-GML data model. These recent ventures are progressing a movement that extends far beyond the 1990s concepts of "smart cities" enabled by wireless telecommunications. In the Digital Earth science paradigm, cities must simulate their key situations and scenarios and analyze Earth observation data obtained via satellite-enabled devices that remotely detect and interpret all the light and radio waves of the electromagnetic spectrum.

Keywords Data cities · Geospatial · Digital urbanism · GEOSS · Digital earth · Earth observations · Smart cities · Urban modeling · Geodesign

16.1 Introduction: Satellites Meet Cities

The Digital Earth project (Gore 1992, 1999; Goodchild et al. 2012; Craglia et al. 2012; Jackson and Simpson 2012) is aligned with the intergovernmental program for a Global Earth Observation System of Systems (GEOSS was launched in 2005, the same year as the online commercial globe Google Earth; Group on Earth Observations (GEO) 2007, 2015; Jackson and Simpson 2012). These and many other satellite

D. Jackson (✉)
School of Architecture, University of Kent, Canterbury, UK
e-mail: dj5@kent.ac.uk

R. Simpson
Department of Infrastructure Engineering, University of Melbourne, Melbourne, Australia

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3_16

and semiconductor-enabled cartography advances imply the need to produce a globally useful schema for more scientific and eco-ethical management of cities (Ratti and Claudel 2016). This aspect of Gore's Digital Earth dream remains far from reality and was promoted earlier by Richard Buckminster Fuller, beginning with his 4D Air-Ocean World Town Plan concept diagram (Fuller 1928; Fig. 16.1), followed by various urban synergetics proposals and prototypes. These contributed to his influential late-career book *Operating Manual for Spaceship Earth* (Fuller 1969), which was published exactly fifty years before this *Manual of Digital Earth*.

How should we design an internationally cohesive and locally effective system for understanding and managing urban stocks and flows around our planet? This question requires comparisons and integrations of significant concepts published and prototyped by key scientists, technology innovators, architects and other leaders of the urban informatics revolution; especially since Fuller died in 1983.

His original World Town Plan sketch was invented when "computers" were mathematically minded people, more than a decade before German engineer Konrad Zuse invented the first electromechanical, stored-program computing machine; his Model Z3 was first demonstrated in 1941. Fuller expired shortly after *Time* magazine named "The Computer" instead of a human recipient for its annual "Man of the Year Award" cover feature (Brosan and Segal 1982).

Fig. 16.1 Fuller's air-ocean world town plan diagram, 1927–28 (Estate of R. Buckminster Fuller/John Ferry)

Although his vision of an electronic infrastructure to operate Spaceship Earth was inspired by radar and airplane autopilot systems, satellite navigation was not commercialized widely for terrestrial vehicles until the early 1990s. Accompanying the advent of GPS (global positioning system) devices linked to American NAVSTAR satellites were magazine and newspaper reports forecasting commercialization of the internet as a “new information superhighway” and revolutionary television and telephony advances (Negroponte 1993, 1995; Gates 1995). Leading professors of town planning and architecture expected computers to accelerate “smart cities” (Gibson et al. 1992) and the MIT Smart Cities Lab was founded by William J. Mitchell in 2003. Other urban prophecies included “fractal cities” (Batty and Longley 1994), the “city of bits” (Mitchell 1995) and “intelligent environments” (Droege 1997). At the time of writing this chapter, the world’s main satellite navigation systems were GPS (US), BeiDou (China), Galileo (Europe) and GLONASS (Russia; Hunter and Hartcher 2019). We suggest that all of these 1990s terms emerged in response to global commercialization of wireless and mobile telecom infrastructure—and that this century’s Digital Earth and GEOSS planetary systems simulations vision demands a new emphasis on the cruciality of satellite-enabled remote sensing data; thus, we now use the term Data Cities when considering the urban aspects of Digital Earth.

All of those end-of-century writers (and others before and since) highlighted that “wireless” (actually extensively cabled) telecom technology was unlocking a crucial new way to understand cities: not as static compositions of buildings and streets but as dynamic, unpredictable and increasingly networked flows of activity and connections. However, until recently (Jackson 2018) there was little emphasis on how satellites have become essential to what Batty called “a science of cities” (Batty 2005, 3; 2013) and Stephen Wolfram termed “a new kind of science” (Wolfram 2002) that would interpret fractal and cellular automata principles of evolutionary growth and behavior.

Satellites allow for today’s environmental scientists and designers to use increasingly sophisticated machines and programs to monitor and simulate various processes that Jay Forrester termed “urban dynamics” (Forrester 1969). City monitoring and modeling are being accelerated by increasing numbers and constellations of Earth-observing (EO) satellites, including squads of tiny CubeSats carrying miniature remote sensing instruments. These include scanners and sensors to scrutinize atmospheric and ocean conditions for meteorological and marine agencies (producing data that are visualized dramatically for television weather reports). They also include many devices that use all the wavelengths of the electromagnetic spectrum to continually survey the world.

Earth observation methods such as SAR interferometry, GNSS reflectometry (GNSS-R), radar altimetry and lidar sensing are revealing many structures and activities that normally cannot be viewed by humans or have been long obscured. Some dramatic recent examples are digital heritage discoveries and detailed 3D mapping of various ancient cities that were lost for centuries under tropical jungle foliage or catastrophic floods. Specialists in digital archaeology can study early stone carvings under thick coats of dirt and moss, and explore fabled burial grounds, perhaps without touching a spade (Venkataramanan 2014).

For professionals developing and managing contemporary cities, satellite-enabled land surveying has become vital to understanding existing circumstances with unprecedented accuracy—allowing designers, decision makers and stakeholders to share the same eyewitness evidence in discussions of proposals and problems.

To understand how satellite technology and data are being applied to solve today's environmental planning and management challenges, Davina Jackson (coauthor of this chapter) devised a matrix diagram of five research themes and their flow-on priorities and projects in government, commerce and public sector contexts. Drafted from 2008 to 2011, it was published in a GEO-sponsored snapshot report on the scope of the GEOSS/Digital Earth project (Jackson and Simpson 2012, 5; Fig. 16.2). All five research themes are being pursued concurrently towards the ideal of a global model of complex environmental systems. They are natural systems modeling (NSM; projects simulating certain area-defined environmental behaviors), building information modeling (BIM; creating virtual models of structures and testing the viability and defects of each design before on-site construction), city information modeling (CIM; 3D mapping, satellite and aerial imagery, remote sensing and data analytics at scales from street corners to megalopolises), virtual nations and networks (VNN; data management and mapping the environmental conditions of countries,

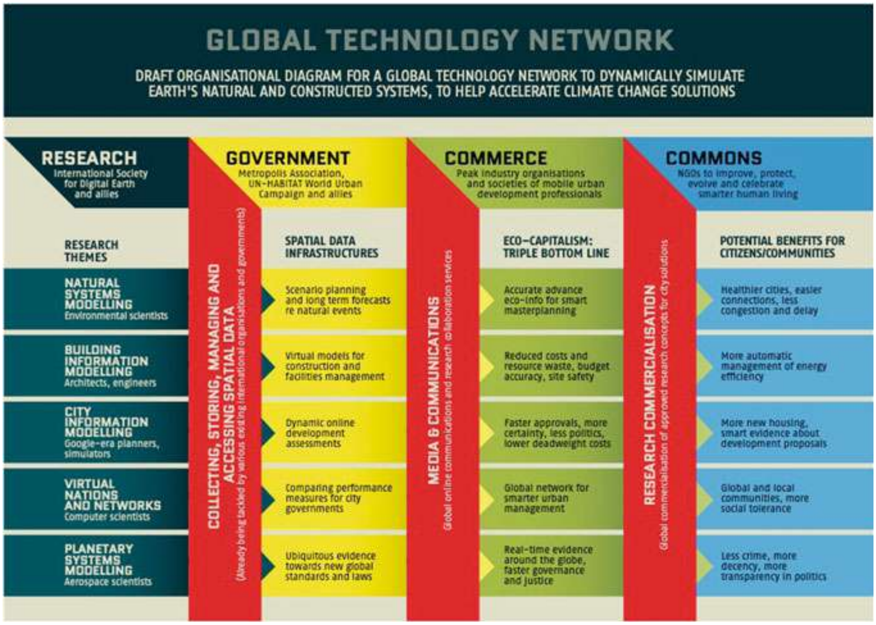


Fig. 16.2 Jackson's global modeling network concept diagram, 2008–2011 (Jackson and Simpson 2012)

multinational regions or continents—e.g., Virtual Singapore) and planetary systems modeling (PSM; integrated 3D data mapping of environmental conditions around the Earth).

All five scales of Earth observations and simulations must be integrated to achieve the concept of a global EO system of systems but the diagram identified them separately to reflect the reality that most researchers operate within specific professional disciplines and domains of study (Fuller 1980). The following sections explain current activities and recent projects that contribute to the integration of data and modeling that is transforming urban development.

16.2 Global and Dynamic Data Mapping of Cities: A New Cartography Paradigm

The Digital Earth vision and GEOSS program are both evolving through collaborations between several hundred international governments, space agencies, science, research and standards organizations, and United Nations entities (UN Global Marketplace n.d.). These groups are organizing different advances towards the system of systems that has begun to allow users to access, analyze, visualize and exploit the data collected by Earth observation instruments aboard or networked with satellites. In this section, we identify how this system is being progressed in ways that may help reform obsolete, insular and ecology-damaging practices by millions of influential actors in urban development and city governance roles.

Today's collaborations are underpinned by shared understandings of the imperative to scientifically tackle the deadly impacts of climate change, including extreme weather events (natural catastrophes), loss of biodiversity, rising sea levels, and extreme heat and drought (UNDP n.d.). Two key UN bodies are leading the task of broadly communicating information and strategies to deal with these wicked problems: the Intergovernmental Panel on Climate Change (IPCC), which releases five-yearly scientific reports recording the world's environmental threats and performance, and the UN Framework Convention on Climate Change (UNFCCC), which organizes annual conventions where relevant organizations discuss, and participating governments agree, how they will reduce ecology-damaging practices in their countries.

Another UN organization, the UN Development Programme (UNDP), globally promotes seventeen sustainable development goals (SDGs) that include climate action and sustainable cities and communities (UNDP n.d.) and other, mainly urban, agendas. Its urban targets for 2030 include upgrading slums, increasing the resilience of communities that are vulnerable to disasters, reducing the environmental impacts of cities, improving air quality and waste management and providing affordable, useful public transport and housing.

Certain places are evidently dangerous to occupy—flood plains, fire and earthquake zones, or countries prone to war. Vast land areas, especially deserts or polar regions, are shunned because they seem inhospitably dry or frozen. Should they be invigorated via hydro- and geo-engineering? This matter is being debated by ambitious scientists and engineers through their academic and professional organizations.

Many properties near water—clifftop mansions on New York's Long Island or entire island states such as Tokelau and Nauru—risk subsidence and submersion through the same forces (freak waves, storms and rising sea levels) that caused ancient monumental cities such as Thonis (Egypt) and Harrapan (India) to slide into the sea. Cyclones sometimes destroy popular resorts on South Pacific islands and towns along Asian coasts, and fires seasonally burn through leafy suburbs in southeast Australia and southern California. Residents of large hillside cities in Central America and South America—like La Paz in Bolivia—understand that their homes suddenly might slip down their slopes of clay. All these dangers appear to be escalating with the global warming that Swedish scientist Svante Arrhenius first predicted in 1896. He calculated that global temperatures would rise by 5 °C with the doubling of carbon dioxide burned into the atmosphere. This prediction seems consistent with today's UN forecasting of a 5 °C temperature increase globally by 2050 (UNDP n.d.).

In Geneva, the International Centre for Earth Simulations (ICES Foundation) archives scientific and press reports of environmental disasters on its website (Bishop 2018). Its articles from September 2017 to March 2018 included photographs of a hotel tower falling after a Taiwan earthquake, bridges collapsing in Colombia and Florida, homes buried under mudslides in southern California, a volcano erupting in Bali and Hurricane Irma battering Caribbean countries and Miami. ICES, led by Bob Bishop, a scientist expert in high-performance, real-time computer simulations, aims to establish an advanced computing facility in Switzerland for modeling complex environmental systems. His foundation offers a worldwide QLARM message service that promptly predicts and maps likely building damage and human casualties after earthquakes (Wyss n.d.). More than 1000 alerts have been issued since 2002 using geological data from the Swiss Seismological Service, settlement records from the *World Housing Encyclopedia* and population statistics. ICES also completed earthquake vulnerability studies of Haiti (for the Swiss Department of Foreign Affairs) and Kyrgyzstan (for Médecins Sans Frontièrs).

In recent lectures, Bishop analyzed the challenges and potential for building "an open, integrated, wholistic model of Planet Earth for decision support, disaster reduction and public good"—the same vision as the GEOSS, Gore's 1990s Digital Earth and Fuller's 1960s Spaceship Earth. He warned that quantities of data—mostly unstructured data—are growing far faster than global computing power—and that both are insufficient to crunch solutions for the world's many serious environmental and sociopolitical threats. He predicted that, as well as quantum computing systems, global simulations projects will ultimately be improved by neuromorphic computing to imitate information processing by human brains. Recent projects to develop neuromorphic chips include TrueNorth by IBM, SpiNNaker by Manchester University and BrainScaleS, initiated by Heidelberg University. The two academic ventures have transferred to Europe's Human Brain Project (HBP), which aims to substantively upgrade today's energy-guzzling computers built with von Neumann architectures (Modha n.d.). Bishop (a former chair of the HBP advisory panel) suggested that, while these computing capacity ambitions are being pursued, current projects can reduce energy consumption by installing processor-in-memory (PIM) technologies to use big data, preferably without moving it. While computer architectures would be global, the data should be applied to solve local urban and regional problems.

European scientists recently published the first sophisticated satellite mapping of the world's human settlements, depicting four decades of population statistics and machine-analyzed Landsat data, including some information on building heights, footprints and materials. Launched in 2012 as the Global Human Settlement Layer (GHSL dataset for GEOSS; Pesaresi et al. 2016), this project is included in the GEO Human Planet initiative that was announced at the 2016 United Nations Habitat III conference on settlements. As well as identifying several major new cities in Asia that were not UN-recorded, it applied astrospatial (developed for space exploration), geomatics (terrestrial monitoring) and telematics systems to the formerly paper-centric domains of land surveying, cartography, architecture and town planning. Outstanding 3D and 4D visualizations of the GHSL data, depicted as "population mountains", were produced by Alasdair Rae (Rae 2016, 2018) and Matt Daniels (Daniels 2018a, b).

The first example of global 3D video-mapping of urban population (including growth) statistics was Japan's PopulouSCAPE project, which included a 10-minute *Night Flight Over an Urbanizing World*, providing aerial views of cities as surging towers of population (visualizing UN Figures) and intercity transport and communications connections (Ito et al. 2005; Team PopulouSCAPE 2005). Another historically significant example of planet-scale modeling of cities was the *Pulse of the Planet* real-time video visualization of AT&T data recording telephone and internet traffic between New York and other cities. Produced by a Carlo Ratti-led team at MIT's SENSEable City Lab, it was shown in the *Design and Elastic Mind* exhibition at New York's Museum of Modern Art (MIT SENSEable City Lab 2008). These two projects were perhaps history's first world-scale depictions of the data cities movement (Jackson 2008; Jackson and Simpson 2012)—following some important mid-2000s video simulations of specific cities such as the *Virtual London* model (showing flood and shadow simulations) by University College London's Centre for Advanced Spatial Analysis (2002–2005) and the *Real-Time Rome* mobile phone data-mapping show at the Venice Biennale by MIT's SENSEable Cities Lab (2006).

One notable new world urban mapping project is the Global Urban Footprint (GUF), led by Thomas Esch's team at the Earth Observation Center of the German Space Agency (DLR, Fig. 16.3). Scatters of tiny black dots show settlement patterns with unprecedented detail and precision, using radar data from the TerraSAR-X and TanDEM-X pair of satellites operated by the DLR and Airbus Space and Defence. Although only depicted in 2D, the GUF shows the global distribution of human settlements with an unprecedented spatial resolution of 0.4 arcsec (~12 m), using 180,000 satellite scenes expressed in grayscale: black dots for urban areas, white for land and gray for water (DLR n.d.; Fig. 16.4). This instantly informative data visualization (seen on-screen via a swirling sphere) refreshes the adage that new technology

Fig. 16.3 Global Urban Footprint map by the German aerospace center (DLR/Thomas Esch)

Fig. 16.4 Sunlight control modeling for Auckland city (1988), with the operative envelopes visualized as "stained glass" windows (Cadabra)

paradigms (such as 3D time-series video-mapping) are not the only, or always the best, tool for communicating specific information in certain circumstances. As an obvious example, audio remains preferable to video when people are walking or driving cars.

16.3 Global Advances in Computer Design, Analysis and Construction

This section highlights five significant advances in data modeling solutions for major challenges in designing and managing built environments.

16.3.1 Environmental Performance Control Envelopes

New Zealand's Resource Management Act, passed in 1991, was ground-breaking legislation for the nation to conserve and sustainably manage its natural resources: land, air and water. It was underpinned by one of the world's first cases of using architectural computer modeling (then known as CAD, computer-aided drawing, now updated as BIM, building information modeling) to pretest the potential environmental impacts of new building proposals. Approval by a New Zealand court of law for uses of 3D computer models as evidence was first granted for the appeal of a planning decision that delayed construction of Auckland's Sky Tower. As the tallest freestanding structure in the Southern Hemisphere, Sky Tower was a radical departure from the city's conservative urban landscape and would not have been publicly acceptable without using computer simulation and 3D visualization to articulate the regulatory, design and environmental impacts (Fig. 16.4).

Before the Resource Management Act, basic prescriptive rules were used by NZ planning authorities to maintain unimpeded sunlight for specific open spaces. This approach was refined during the years after the 1987 stock market crash, when city property values slumped. In 1988, the Auckland City Council commissioned Cadabra, an applied computer graphics consultancy led by Richard Simpson (coauthor of this chapter) to develop one of the world's first performance-based 3D virtual city models to allow for patterns of sunlight to be more specifically and accurately simulated.

Cadabra's approach was to calculate an overarching operative control envelope (OCE) for the central city. This performance-based envelope was generated from an accurate 3D terrain model of the city, including twenty-seven designated public places (mostly parks) that were to maintain access to sunlight and views. The sunlight controls for these places were evaluated for every moment of the year to determine the overall impact that each would have on the city height limits. The result was a set of twenty-seven envelopes that intersected in complex and dramatic ways in 3D and 4D space above the virtual city skyline. The overarching operative envelope was determined by overlaying all the place-specific envelopes and defining the combined surface minimum to define the OCE.

This operative surface had the appearance of an exaggerated rugged terrain in the sky. It defined a "battle" between ground-based controls. In one place, one control might have a sweeping influence on the height of a potential new building, then would be overridden by another control. The hilly nature of the inner city, protection criteria, and eclectic formats of park boundaries all contributed to this complex expression through controls on the generated operative envelope. Visually, this model appeared like a paint-by-number on a wildly rumpled canvas.

The operative envelope was visualized as a contour plot of heights above a datum. The model enabled performance-based sculpting of the city's urban envelope and regeneration of the individual sunlight controls to ensure solar irradiation for any place for specific periods of the year and times of day. It was also rendered in 3D with proposed and existing 3D CAD models of buildings. If a building complied with the control, it would be visually obvious as it would not breach the envelope. The rendering treated each control as a differently colored "stained glass" window and thereby visualized the volumetric influences of controls in the airspace and throughout any day. Colored light for a specific control might flood the ground to clearly show the influence of any specific control through a day or year.

The final design of Auckland's OCE was published as a set of regulatory contour maps in the 1991 district scheme. This work defined the aesthetic balance and shape of Auckland's skyline and enabled a paradigm shift from obsolete prescriptive controls to more evidence-precise and context-responsive performance controls. The modeling removed legal ambiguities and provided more clarity and certainty for citizens to enjoy maximum sunlight and views when using public spaces.

16.3.2 Geodesign

In 2001, Pascal Mueller, then a postgraduate student at ETH-Zurich's Future Cities Lab, introduced CityEngine, a procedural modeling program to rapidly generate and modify basic forms of buildings in urban scenes. It offered designers flexibility to change the heights and floorplates of specific buildings on the fly, using process scripts rather than the prescriptions of parametric modeling. When commercially launched in 2008, it could generate a realistic 3D online (flythrough) depiction of Venice's Gothic-Byzantine building stock in a few minutes. Users could transform the heights and areas of one or various buildings, and simulate shadows cast at different times of day and year. Since 2011, CityEngine has been owned and updated by Esri, the world's largest commercial GIS mapping software supplier, to integrate with its formerly offline and 2D (pre-Google Earth) suite of ArcGIS mapping tools. Esri's transition to online 3D dynamic modeling and what-if design tools, and its SYMAP-derived topographical data mapping, catalyzed a new educational-promotional venture that the company's founder, Jack Dangermond, launched with the name Geodesign in 2010 (Steinitz 2012, 2013; De Monchaux 2016).

Until recently, the GIS packages needed by surveyors and land-planning professionals have been used separately from the building design (BIM and CAD) programs required by architects and building engineers. However, some surveying firms such as Britain's Severn Partnership are marketing "Scan2BIM" skills to provide point clouds of 3D laser-scanned, geotagged data about existing buildings and landscapes, giving precise bases for modeling structural alterations. Another example is the *ThermalMapper* project in Bremen, where Dorrit Borrmann and Andreas Nüchter from Jacobs University recorded a 360° aerial point cloud of the city square with temperature data from nine thermal images overlaid on eleven 40° laser-scanned poses of the historic buildings, plaza and streets. They used a Riegl VZ-400 3D laser scanner and an Optris PI infrared camera, with the pose files calculated using 3DTK based on odometry information (Borrmann and Nüchter n.d., Fig. 16.5).

In 2017, Autodesk and Esri announced a new collaboration to integrate their building design and environmental data modeling tools. Patrick Janssen, with the Singapore ETH Future Cities Lab, was skeptical that this partnership would resolve all

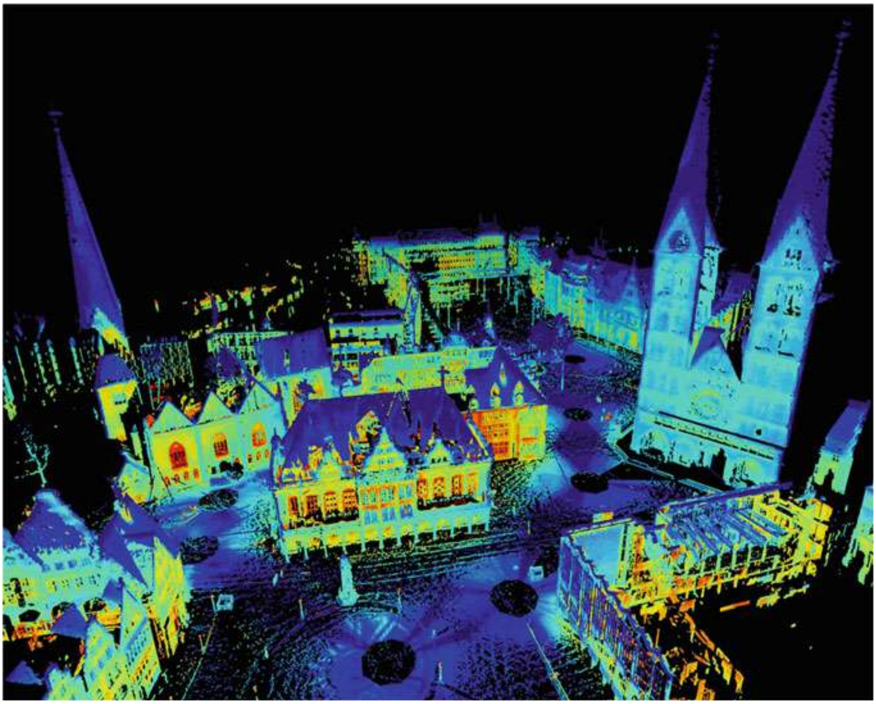


Fig. 16.5 *ThermoRathaus*, a point cloud of Bremen's city square, overlaid with temperature data from a thermal imaging camera (Borrmann and Nüchter, Jacobs University)

the aspirations of architects but recognized opportunities to resolve significant deficiencies in Esri's GIS and Autodesk's BIM and graphics packages. Janssen praised GIS software for providing data online and downloadable in tiles and criticized BIM programs for not providing online access to data and fine-grain imagery. He said advanced architects were eager for open semantics, "where you can compute your own stuff"; a capability now available to some extent in Houdini but likely to be more prevalent soon with new developers using Amazon Web Services or Web Assembly (Wasm). He said that most sophisticated modeling programs remain too complex in their structural routines—generating too many links and nodes in their data graphics. He was developing a new online geometrical design application, Vidamo, which he expected to be more efficient and more comprehensible for users without advanced programming skills (Jackson 2018, 37).

At the 2018 Geodesign Summit, Dangermond's presentation emphasized "the science of where" as being crucial for "understanding and managing our world" through "integrating people, processes, things and data about them" via three types of information infrastructure: records, insights and engagement. He also highlighted three major groups of trends that would evolve WebGIS during coming decades (and which are applicable to other types of software needed for urban planning and design). For professionals concerned with *data*, relevant advances include drones, lidar, scientific measurements, real-time video, crowdsourcing and much more detailed information on traffic, demographics, weather and locations. For experts developing *computer infrastructure*, he highlighted mobile communications, big data, machine learning, distributed computing, SaaS (software as a service), the IoT (Internet of Things), cloud storage and parallel computing, web services, microservices and networks. For *GIS innovation*, he focused on "expanding the power" via advanced analytics, open APIs, dynamic image processing, online content, apps, 3D modeling and smart mapping, data exploration, hubs, real-time visualization, Python programming and portals (Dangermond 2018).

As a corporation led by landscape architecture graduates from Harvard, Esri is focused on how to use computer tools to eco-sensitively integrate buildings and urban infrastructure with natural environments. Echoing three of the five research themes identified in Jackson's 2007 GEOSS-DE network diagram (Jackson and Simpson 2012, 5; Fig. 16.2), Dangermond identified four main types of modeling that should be increasingly integrated: landscape information models (LIMs; another term for natural systems modeling), building information models (BIMs), city information models (CIMs) and zoning information models (ZIMs; a CIM subset of particular value to government planners; Dangermond 2018).

16.3.3 Digital Engineering and Digital Twinning Standards

Also called virtual engineering, digital engineering is a shorthand reference to the consistent use of digital methods and tools throughout product development and production processes to improve planning quality and process controls over an asset's entire life cycle. Digital twinning involves modeling and simultaneously sustaining all the virtual systems associated with a physical entity. These terms emerged with computer modeling and autopiloting systems developed for aerospace, ship and car manufacturing and operations. Now these labels have transferred to BIM and geospatial environmental modeling—which evolved from different 2D paradigms (CAD drawing and GIS mapping) that depend on different methods of structuring data.

To establish interoperability, the buildingSMART International (bSI) group and the Open Geospatial Consortium (OGC) established a joint working group in 2017 to prepare a roadmap towards a global standards framework named the Integrated Digital Built Environment (IDBE). BSI administers Open BIM standards, including Industry Foundation Classes (IFCs), and the OGC administers OpenGIS, including geospatial data interoperability, and the Reference Model and GML standards. The IDBE is intended to underpin digital engineering and enable digital twinning of physical conditions with corresponding records held in digital repositories. The physical twin may be represented in the digital twin (virtual model) at any level of detail (LOD). However, there are new moves beyond formerly prescriptive notions of LOD to a more agile, performance-based, level of information needed (LOIN) approach, which specifies why data are required, what specific data are required, when they are required, and who is responsible for the transfers and uses.

16.3.4 Astrospatial Architecture

Architecture's ancient history switched tones around the turn of the third millennium. In May 2000, Aaron Betsky published *Architecture Must Burn*, a critique of late-twentieth century architectural culture and a "manifesto for architecture beyond building" (Betsky and Adigard 2000). This book preceded, by just eighteen months, the explosions that collapsed America's twin towers of modern capitalism, the World Trade Center, in September 2011. Five months later, Manhattan architects Diller Scofidio+Renfro revealed an unprecedented anti-icon: the Blur pavilion, a wide cloud of clean water vapor hovering low across Lake Neuchâtel during the 2002 Swiss National Expo. Solar-powered and with sensors dotted across its fog-obscured steelwork, this work symbolized two novel impulses: to evaporate architecture's antiquated focus on merely crafting static structures using weighty materials dug from the Earth and to steam-clean a world remaining stubbornly reliant on carbon-belching fossil fuels.

Blur catalyzed a post-internet design movement that was later named astrospatial architecture (Jackson 2016). Protagonists now intend to design extraordinary compositions of solids and voids and devise memorable interpretations and experiences via light and data. For example, in 2011 Joseph Paradiso's Responsive Environments group at MIT's Media Lab revealed perhaps the world's first video simulation of invisible atmospheric dynamics inside a building—using temperature, humidity, light, sound, human movement and other data, streaming from RFID and other sensors around floors of the real building. Visualized with the Unity game engine, the *DoppelLab* showed how space pulsates with unseen information (Paradiso 2011).

Architectural technology has evolved from computer-aided drawing (CAD) tools (beginning with light pens drawn on tiny screens in the late 1950s) to BIMs that are derived from aerospace engineering software and can script algorithms to operate fabrication machinery such as 3D printers and construction robots. Professors leading this international revolution tend to be involved in three overlapping design movements: Parametricism (Schumacher 2008), Smartgeometry (Smartgeometry n.d.; Peters and Peters 2013) and the more recent Advanced Architectural Geometry (AAG) group (Adriaenssens et al. 2016).

Building models created in programs such as Autodesk's BIM360 or Revit, or Trimble's Tekla BIMsight or Connect, can be exported for viewing with headsets using plug-ins such as Modelo, Prospect, Enscape, Umbra or AUGmentecture. Another capability was demonstrated by Greg LynnFORM at the 2016 Biennale of Architecture in Venice, where Lynn and his team used Microsoft HoloLens goggles and augmented reality (AR) software to compare multiple holographic scale models of the Tate Modern building in London with a physical scale model of a giant former car plant in Detroit (Fig. 16.6). HoloLens wearers could look inside the physical model and walk around full-scale virtual rooms defined by lines of ephemeral light. The team also highlighted and overlaid the history of the building being redesigned and showed different road and aerial vehicles flowing around the site (Jackson 2018, 19–20).

Beyond the design studio, VR, AR and nonimmersive 360° viewing systems are valuable tools for the property industry (Stanley 2017). They help clarify building proposals to stakeholders influencing council development approvals. They help in marketing buildings and apartments prior to completion or to inform remotely located investors. For example, spherical imagery captured by drones can clarify views from different floor levels of an unbuilt tower. The industry expects continuous

Fig. 16.6 Greg Lynn finger-snaps holographic (AR) building models of London's Tate Modern onto his physical model for car plant redevelopment in Detroit (Microsoft)

improvements to VR and AR experience kits and 360° cameras over the coming decade. Better graphics (pixel densities and spatial resolutions), audio, haptics (more touch-sensitive handset interfaces), and tracking speeds are still needed.

Another way to represent architecture and cities is with holographic images. These are data-coded recordings of light fields (waves of particles scattered off illuminated objects). Like sound recordings, holographic recordings can be reproduced later—but usually with considerable loss of fidelity, so the representations seem ethereal. Virtual building and city models can also be converted into 3D holographic imagery printed on film. When illuminated from above, these renditions seem to pop in three dimensions from their glossy sheets. If not precisely lit and viewed, holographic images appear spectral and chromatically fragmented.

Holographic glasses and headsets underpin augmented reality (AR)—a domain alternatively named "metasensory augmentation" by wearable computing visionary Steve Mann. At MIT in 1978, he prototyped the first AR spectacles, Digital Eye Glass, and later versions could be finger-tapped to convey holographic data. Those precedents inspired Google's Glass smart specs (sold generally from 2013 to 2015), Microsoft's HoloLens system (launched 2016) and the Vuzix Blade smart glasses (previewed in 2018, Statt 2018). The HoloLens has been surpassed technically by another of Mann's creations, Metavision's Meta 2 visor, released to Unity game developers in late 2017.

LiFi (light fidelity) is another emerging technology expected to energize built environments. Demonstrated by Harald Haas in a TED talk in 2011, a LiFi system uses the semiconductors of LED lamps (such as downlights) to transmit data (Haas 2015). In some early tests, LiFi networks transferred data at much faster speeds than is currently possible over WiFi networks conducting low-frequency radio waves and microwaves. This is because LiFi uses the higher frequencies and bandwidths that come from the visible light, infrared and near-ultraviolet waves that share the mid-range of the electromagnetic spectrum. Since Haas and his partner Mostafa Algani set up the pureLiFi company to commercialize his discoveries, several dozen start-ups and corporations have begun developing LiFi applications using next-generation LEDs with signal processing capabilities. In Dubai, Zero 1 used LED streetlights for networking data—exploiting pre-Haas research on urban transport-logistics telematics. In Dresden, the Fraunhofer IPMS research center has developed industrial automation solutions for several corporations. All major electronics manufacturers—General Electric, Panasonic, Samsung, Philips, Osram, Qualcomm and Cree—are racing to market with LiFi data-and-light product suites.

People in polar countries often feel depressed by the long nights of winter—needing treatment with mood-elevating colors and wavelengths of light. This affliction, called seasonal affective disorder (SAD), seems to emerge from changes to a body's circadian rhythm and serotonin and melatonin hormone levels. Some local governments in near-Arctic latitudes encourage their citizens to take therapy sessions and cheer their communities with winter light festivals (that also magnetise tourism). From 2012 to 2016, Oslo artists Christine Istad and Lisa Pacini responded to SAD in Norwegian towns (where there is no daylight during January) by trucking around a night sun—a 3 m-diameter circular panel crusted with hundreds of color-changing

LEDs (Anon 2013). A popular, three-dimensional, night sun is Rafael Lozano-Hemmer's *Solar Equation* aerostat (helium balloon), which is video-mapped with layers of computer-generated imagery and data derived from NASA's sun-monitoring instruments (Fig. 16.7). These are just two examples of creative urban (outdoor) applications advancing this century's revolution in 'electroluminescent' technology, based on semi-conductor controls of electric pulses (Neumann and Champa 2002; Jackson 2015).

16.3.5 Artificial Intelligence

Artificial intelligence was ignored by most built environment professionals until the internet caused widespread apprehensions during the 1990s, systemic disruption during the 2000s, and now, inevitably, new ways of understanding and doing things. Today, AI brings another wave of unfamiliar technologies and terms—including augmented intelligence, where machines are intended to improve human abilities to decide and perform. This seems less threatening than artificial intelligence, where machines are presumed to increasingly replace humans to a tipping point that Ray Kurzweil termed the Singularity (Kurzweil 2005).

Fig. 16.7 *Solar Equation,* a "night-sun" (helium balloon) designed by Rafael Lozano-Hemmer and mapped with NASA sun-monitoring data (Marcel Aucar)

All intelligence, artificial or natural, flows from competent processing of information. Most AI researchers have abandoned their early reliance on preprogrammed rules to solve problems. Instead, they are evolving machine learning, where computers use statistical learning algorithms to gradually teach themselves how to intelligently process big data. The more data that computers are given, the more capably they can perform complex tasks; partly through their supra-human powers of pattern recognition (*New Scientist* 2017).

Dacheng Tao at the University of Sydney has developed a classification system to help understand different concepts, methods and challenges that are being advanced in AI and robotics. His taxonomy highlights four basic functions performed by machines: perceiving, learning, reasoning and behaving. Devices fueled by data from sensors and cameras must perform one or more of these functions to help solve humanity's ultra-wicked problem of how to sensibly manage our planet.

Machine learning is a new field that is being divided into different specialties: unsupervised learning (training machines to identify untagged images), supervised learning (training using labeled or annotated information), reinforcement learning (training via rewards for correct actions) and deep learning (using complex neural networks). Neural networks are software circuits inspired by flows of information through human brains. They can deliver general artificial intelligence (solving various tasks) or narrow AI (expertise in one or two specific tasks).

One of the most promising potentials in AI is for robots to replace humans in performing extremely dangerous tasks: such as exploring nuclear power plants after an explosion, entering narrow cavities to replace damaged wiring or recording stress points in unstable structures. Czech writer Karel Čapek first coined the term robot in his 1921 play *R. U. R: Rossum's Universal Robots*, and today's humanoid versions such as Boston Dynamics' Atlas and Honda's Asimo are agile and realistic. Most of Boston's robots, being developed with the US Defense Advanced Research Projects Agency (DARPA), emulate fleet-footed animals and are intended to replace soldiers on topographically rugged battlefields. Swiveling, fixed-footed robots (mounted on a floor or ceiling) can print small masonry dome structures and assemble timber-framed houses (Jackson 2018, 42–48; Kohler et al. 2014).

Researchers developing computer vision systems are evolving improved ways for cameras, sensors and software to detect, recognize and track moving objects, including people, analyze environments by segmenting items of interest in changing scenes, estimate distances between cameras and objects in view, and enhance the clarity of images. Face-detection software can discern and frame almost every head in crowds of thousands. Any newly scanned face can be matched instantly with the same face from a digital archive. Data analysts can also clarify blurry, hazy, too-dark, wrongly colored and low-resolution images using smarter versions of the photoenhance tools found on standard laptops and smartphones. As these perception technologies improve, CCTV is becoming all-pervasive, with predictable reductions of both public crime and personal privacy.

Machine vision scientists depend on open-source datasets comprising images of objects that are classified and labeled to allow for comparisons with new images containing similar objects. The world's largest object dataset, ImageNet, contains more than 14 million crowd-labeled thumbnails, which can be downloaded to help identify different types of natural places, buildings, rooms, products such as fridges or dishwashers, furniture, fabrics, clothes, and apparel such as hats or sunglasses. Vision experts classify database images according to whether they depict "things" (box-frameable objects such as chairs, people or windows) or "stuff" (matter with no clear boundaries such as a patch of sky, an office corridor, a wall, a hillside or a street) (Stanford Vision Lab 2016).

Ironically, the image databases now being assembled to support AI analytics all depend on the "artificial artificial intelligence" (i.e., the non-electronic knowledge of humans working online) to label and cross-check the images uploaded by database compilers. One busy conduit is Amazon's Mechanical Turk (AMT) portal, which matches employers (such as public research groups) with freelancers to contribute to specific human intelligence tasks (HITs). One recent HIT, to assemble and correctly label 328,000 thumbnail images of "common objects in context" for the Microsoft COCO dataset, required 70,000 h of work by Microsoft-funded AMT participants (Lin et al. 2015).

Cameras and scanners capture images that can be analyzed, compared and manipulated, increasingly automatically and accurately. Some powerful processes are becoming common practices for owners and managers of major buildings and public places. For example, different faces, facial expressions, poses and walking gaits can be transferred or morphed between source videos featuring one or more people and "target" videos involving other people. Security cameras can detect licence plates and simultaneously track clusters of moving vehicles, even at night. All these observation systems are being integrated gradually with traffic lights, smartpoles and building heating, ventilation and cooling (HVAC) systems.

Data networks underpin the sensing and imaging infrastructure that is necessary to deliver key goals for the Global Earth Observation System of Systems. These include improving urban and disaster resilience; public health; energy, mineral and water resource management; infrastructure and transport systems; food security and agriculture, and healthy, biodiverse ecosystems. These domains overlap—integration of information and technology solutions is the main point of the GEOSS.

Caution pervaded a recent editorial for *Environment and Planning B*, in which Michael Batty warned readers not to expect too much sophisticated intelligence from "intelligent" machines. He said that machine learning through highly repetitive schemes of pattern recognition is "not much more than sophisticated averaging" but because machines could rapidly process vast quantities of data, they would continue to be useful for automated tasks such as monitoring and prediction of energy uses, delivery of location-based services, and transport. He suggested that machines would not be capable of replacing humans in planning long-term development of cities because they could not compute "the hard choices" of how a city functions economically and is organized in terms of social equity. He suggested both exploration of the limits of AI in understanding cities and "a concerted effort" by planners to invent new ways of automating urban functions (Batty 2018).

16.4 Some Recent Urban and Regional Case Studies with Global Potential

This section highlights seven projects and research groups that have advanced urban environmental simulations with methods that could be applied in many other cases and places.

16.4.1 MIT Media Lab Projects, United States

Carlo Ratti's SENSEable City Lab worked with government leaders in Cambridge, Massachusetts to prototype the *City Scanner* project to acquire weather and air quality data for different precincts using sensors fixed to garbage trucks. Christoph Reinhart's Sustainable Design Lab developed an urban modeling interface (*umi*) program that evaluates key environmental performances of neighborhoods and cities. First, the area being studied is architecturally modeled in Rhino 3D, and then the model is analyzed for walkability, daylighting and several types of energy consumption (Fig. 16.8).

Kent Larson's City Science group is continuing its *CityScope* project to model city precincts using color-coded Lego bricks, which are sensor-tagged and plotted on screens as data units. Users rapidly move the data bricks to reveal different ways

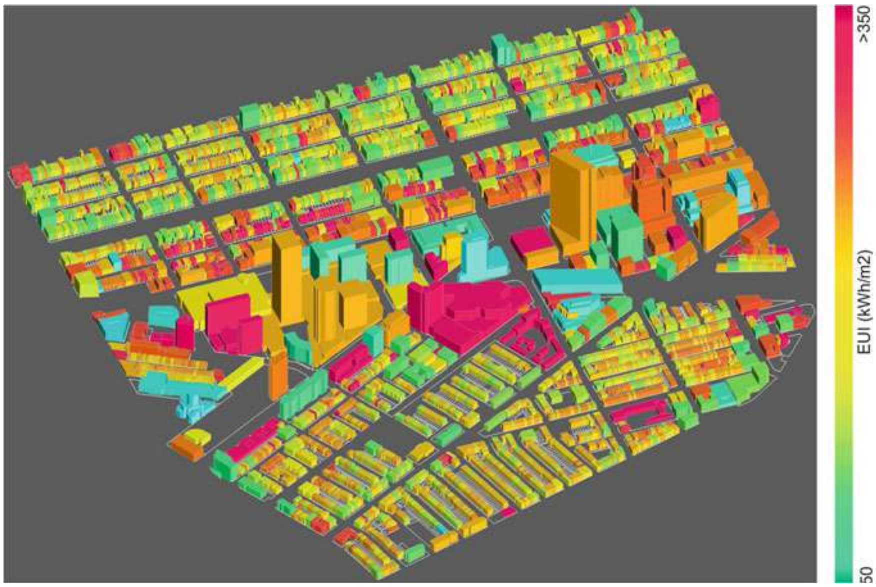


Fig. 16.8 Energy use analysis of a Rhino 3D city model in the urban modeling interface (umi) developed by MIT's sustainable design lab

to improve the density, proximity to services, and demographic diversity (vibrancy) of each area. City Science researchers have also prototyped ingenious "mobility on demand" (MOD) solutions. Their latest persuasive electric vehicle (PEV) can drive autonomously, even following its human passengers slowly if they decide to walk themselves; can be driven in bike lanes without the driver requiring a vehicle license; could be suitable for public sharing and can move both people and goods. In another project, Larson and Hasier Larrea showed how five hundred people could be housed in a medium-rise block of 25 sqm "action apartments" with the same footprint as forty-five conventional car spaces. Since graduation, Larrea has established a company, Ori, to make and sell these robotically mobile furniture suites.

16.4.2 Almere 2030, the Netherlands

In the Netherlands, one of a few countries noted for consistent innovations in urban spatial planning, architects MVRDV (Maas van Rijs de Vries) designed a 2030 vision plan to help the municipality of Almere plan polycentric growth on 250 sq km of polder land reclaimed in the 1960s (Fig. 16.9). The terrain is three meters lower than the water level of the adjacent IJseelmeer (lake) so Almere is constantly at risk of flooding, protected by a system of dykes and sluices. MVRDV's plan, gradually underway now, contradicted the popular Western strategy of transport-oriented development (TOD), where high-rise apartment buildings are clustered around suburban metro stations and new ribbons of low-to-medium-rise housing and commercial development are encouraged along main bus routes and rail lines.

Fig. 16.9 MVRDV masterplan for four "carpet cities" at Almere, the Netherlands

Instead, MVRDV proposed four "carpet cities". IJ-Land, designed with California architect William McDonough, will be a series of new island nature reserves in the lake, including 5–10,000 homes. Pampus will be a high-density, medium-rise village of 20,000 partly floating homes, with all streets leading to a lakefront boulevard. Almere Centre will extend the current city center with development of Almere Floriade, a compact and ultragreen neighborhood intended to be the horticultural campus for the World Expo in 2022. The public arboretum will contain 1,600 new homes, offices and facilities. In Oosterwold (Freeland), the first residents have begun to build their own neighborhood, with up to 15,000 new homes to be set in agricultural fields east of central Almere (MVRDV 1999, 2007, n.d.).

16.4.3 Jade Eco Park, Taiwan, China

French architect Philippe Rahm redesigned an obsolete airport in Taichung to provide a "meteorological" recreation landscape, Jade Eco Park, where vegetation and paths are interrupted by freestanding structures comprising white pipes, air ducts, sensors, filters and other electronic devices (Fig. 16.10). These were designed to mitigate Taichung's generally hot, humid climate and air pollution: they blow cool breezes, release mists or patches of rain, or clean local air to generate three types of artificial and contained atmospheric experiences: Coolia (four cool zones), Clearia (four areas of clean air) and Dria (three areas of dry air). Rahm's team first monitored and mapped the existing temperatures, humidity and air pollution conditions across the site and then used computational fluid dynamics to create an atmospherics map of the site. The

Fig. 16.10 Illustration of clean, cool and dry atmospheres generated across the Jade Eco Park in Taipei, designed by a team led by Philippe Rahm

new atmospheric zones overlap each other to allow for different sensory experiences to be selected at different times of day or during the year (Jackson 2018, 54; Rahm n.d.).

16.4.4 Nocturnal Barcelona, Spain

In Barcelona, the "datatecture" studio 300.000 km/s (speed of light) mapped the city's current and potential night-lighting of streets and public squares in a 2017 report for the city council's Municipal Institute of Urban Landscape and Quality of Life (Fig. 16.11). The project included analyses and visualizations of data on mobility, citizens' activities and business types in each location. The report also included comparison pairs of day and night photographs of city scenes, a satellite image of light pollution and the city's lighting regulations.

Fig. 16.11 Map of night lighting around central Barcelona by 300.000 km/s

16.4.5 Spatial Information Management Platform, Australia

Cities are complex, adaptive systems (CAS), where various dynamic systems evolve in interdependent ways. To "digitally twin" (computer-simulate) a city successfully in the Digital Earth context, this complexity needs to be represented accurately within the digital framework. Each component of the model must be able to evolve independently of the whole. Any viable CAS invariably evolved from a simple system that worked. Any complex system designed from scratch typically fails and cannot be repaired.

Modeling the behaviors of real-world complex systems with counterpart digital systems deepens our knowledge and improves our control of real-world scenarios. Recent projects undertaken by a Brisbane geospatial planning consultancy, Meta Moto, have adopted a CAS framework so that complementary systems of record, engagement, and insight can interact through a common semantic ecosystem supporting master data management and spatial data transformation functions. By adhering to open standards and exchange formats, the complementary systems can be made agnostic to one another. For example, graphics library transmission format (glTF) is a royalty-free specification for the efficient transmission and loading of 3D scenes and models by applications. This format defines the sizes of 3D assets and the runtime processing needed to unpack and use those assets. It defines an extensible, common publishing format for 3D content tools and services that streamlines authoring workflows and enables interoperable use of content across systems of engagement (such as web base viewers). Adopting this as a pipeline within a CAS framework ensures that the uses determine the engagement tools, and the semantic data model drives the user experience and presentation of the data in these tools. This approach removes the risks of vendor lock-in and ensures that the system has continuous opportunities to evolve. Meta Moto recently used glTF for data visualizations supporting Brisbane's Cross River Rail project and the next-generation spatial platform for South East Water in Melbourne (Fig. 16.12).

16.4.6 Greening Greater Bendigo, Australia

Bendigo, a regional city of 100,000 people in central Victoria, Australia, recently began to use EO imagery from Europe's Sentinel satellites to regularly monitor changes of vegetation around its towns and suburbs. A Melbourne landscape consultancy, Office of Other Spaces, analyzed sixteen multispectral Earth observation images (captured at 10 m resolution and stacked as a time-series 'data cube') to monitor seasonal and area-specific changes in vegetation throughout the city during the summer from December 2018 to February 2019. This pilot project allowed the city council to create new a vegetation cover benchmark (named the consolidated

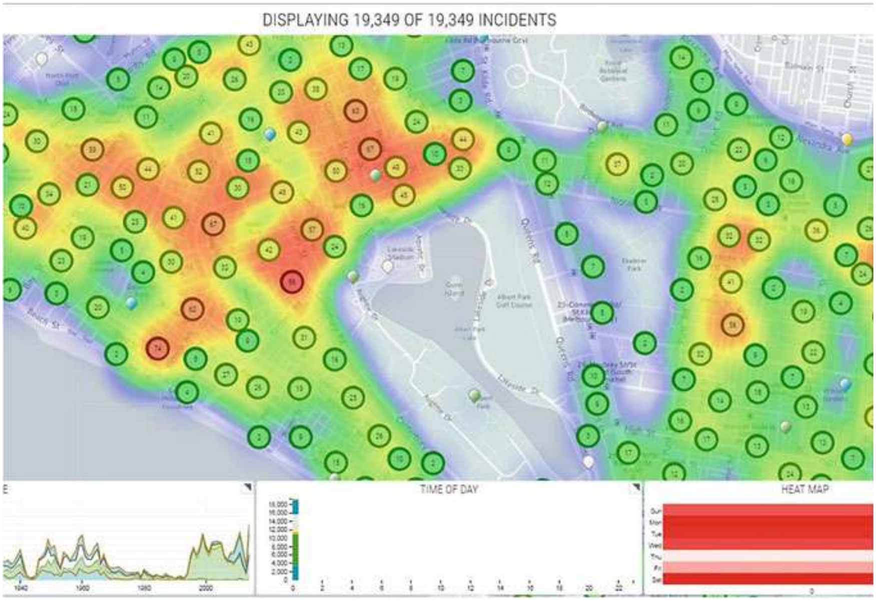


Fig. 16.12 Heat mapping of incidents over time is an example of a system of insight for South East Water, Victoria

mean value, CMV) to clarify the measure of its vegetation cover and determine any site-specific risks of new building developments destroying existing vegetation (Fig. 16.13).

16.4.7 Happy, Smelly and Chatty Maps, Britain

Computer scientists with the Cambridge node of the Nokia-Bell Labs network are advancing a Good City Life program that is intended to support happier citizens—not necessarily the prevailing corporate-governance agenda of 'smart' (time- and cost-efficient) cities (Aiello et al. 2016; Quercia et al. 2016). Led by Daniele Quercia, the Bell-Cambridge social dynamics team has been surveying, analyzing and mapping how people are enjoying—or could better enjoy—the sights, smells, sounds and other atmospheric experiences of public places that they navigate regularly (Fig. 16.14). With other computer scientists at Yahoo Labs in London, and with the Universities of Turin and Sheffield, they analyzed diverse social media tags and developed sophisticated algorithms which allow users of mobile devices to generate navigation routes on their GPS map apps (initially tested via OpenStreetMaps for specific zones of London and Barcelona). The paths they calculate are usually longer than the quickest journeys but can provide happier scenes, smells and sounds along the

Fig. 16.13 Time-stacked EO images of Bendigo, analyzed for urban vegetation changes by Melbourne's Office of Other Spaces

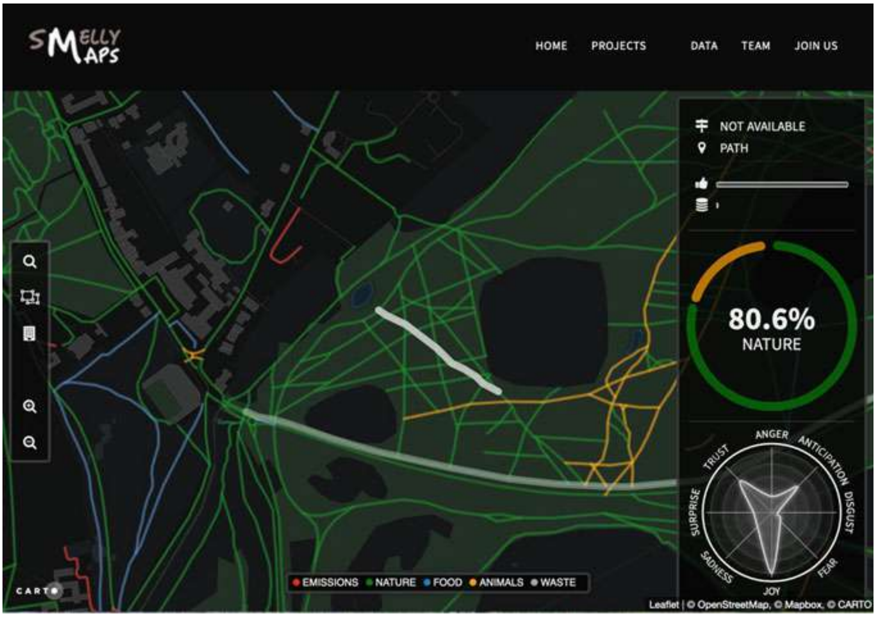


Fig. 16.14 Data-visual street map indicating five types of smells and likely emotions for people navigating an urban area, by Nokia-Bell Labs, Cambridge, UK

way. In recent papers, Quercia's team noted that urban planners and local government officials tend to focus on unpleasant odors and acoustics (signaling air or noise pollution) and mostly ignore desirable smells and sounds.

16.5 Urban Criteria, Process and Standards Taxonomies and Platforms

Many of the world's most intelligent and experienced urban professionals are not digital natives, sometimes refer to themselves as Luddites, and still use hand-inked diagrams, lists, tables and grammatical sentences to express their ideas. Some of these thought-leaders have devised schema for systems that could someday be programmed to analyze development proposals and building performance. This section reviews some taxonomies (matrices of criteria, goals, procedures and standards) that should be integrated into the future GEOSS-Digital Earth world-management system. These develop Fuller's concepts for what he initially named the 4D Air-Ocean World Town Plan then termed the World Game and Spaceship Earth. In one of his last books, Fuller said:

> This design revolution must employ a world-around, satellite-interlinked, data-banks-and-computer-accomplished conversion of present-day, exclusively geocentric, Spaceship Earth wealth accounting [… to a system where …] computers fed with all the relevant energy-efficiency facts will be able to demonstrate which uses will produce the greatest long-term benefit for all humanity (Fuller 1980, 199, 225).

Two critical factors lie at opposite ends of today's project to deliver Fuller's vision. One is the need to decide what data to collect to understand the world's conditions and how to organize it effectively (the main concern of UN-GGIM and other UN technical agencies, the Group on Earth Observations (building the GEOSS Common Infrastructure, GCI), CODATA, the Open Geospatial Consortium (OGC) and the (now disbanded) Global Spatial Data Infrastructure Association (GSDI; Crompvoets et al. 2008). The other challenge is for leaders of international land development organizations to clarify how to identify and deliver all data relevant to future modeling of eco-ethical developments. This agenda is being led by the World Bank (Global Platform for Sustainable Cities), the C40 Large Cities for Climate Change group, the Council on Tall Buildings and the Urban Habitat (CTBUH), the World Federation of Engineering Organizations (WFEO), the World Green Building Council (WGBC) and the OGC-buildingSMART alliance.

Guiding the collection and organization of satellite data are the essential climate variables (ECVs) for all three domains of our planet's environment: atmospheric, hydrographic and terrestrial. The ECV datasets are intended to underpin future eco-ethical practices in land development, and include river discharges, water use, groundwater and lake levels, snow cover, glaciers and ice caps, permafrost and seasonally frozen ground, albedo (surface reflectance), land cover (vegetation types), photosynthetically active radiation, leaf area, biomass, fire disturbance and soil moisture.

Geodata, urban planning and natural resource administration agencies will need to collaborate to obtain the ECV data relevant to their domains of governance and to integrate it with the cadastral and other topographical datasets that they administer. Also valuable are the lidar and radar scanning and 3D city modeling data that are obtained by commercial EO satellite operators (including DigitalGlobe, Airbus and Urthecast), aerial survey companies (such as AAM, Nearmap and Borbas), and locality-diverse providers of terrestrial environmental imaging and remote sensing services.

16.5.1 CityGML and 3D Cadastre

CityGML is an XML-based, 3D vector, open data model for storing and exchanging 3D city and landscape models that is based on the geography markup language (GML) produced by the Open Geospatial Consortium (OGC) and the International Standards Organization (ISO TC211). It defines the objects, properties, aggregations and relations contained in models, allowing for them to be readily compared and for correctly classified data to be reused. The platform allows for sophisticated analysis tasks and thematic inquiries relevant to most urban professions and management functions. CityGML is evolving continuously to improve 3D and 4D city modeling. Dynamic variations of the properties of a city object can be represented using the Dynamizer feature type. This enables specific objects in the 3D model to be linked with simulations or time series data. This can be used to trigger dynamic behaviors such as transformation of the geometry, thematic data, or the appearance of a specific object. This event-driven dynamic sentience of a city model is a foundation for advancing digital twinning of the physical world.

Current urban development approaches require more sophisticated conceptualizations of spatial data and new tools to holistically facilitate four key spatial planning tasks: urban management, impact assessment, site and road selection, and strategic planning. Sabri et al. (2015) developed a new framework to leverage current 3D geospatial and data model technologies in urban modeling and analysis. This framework, including recent insights by Biljecki et al. (2014), adopts a new conceptualization for CityGML that covers most 3D city modeling requirements. Sabri's team demonstrated how complex 3D urban scenarios enable city designers to have a greater understanding of existing and proposed urban forms and potential urban heat islands. The study showed how 3D analysis plays a critical role in examining the impacts of urban consolidation strategies and the densification of inner cities. Nevertheless, the 3D level of detail should be enhanced to support more accurate decision making.

In a recent study, Agius et al. (2018) explained how rule-based 3D city modeling enables planners to measure the physical impacts of building controls (e.g., heights, shadows, setbacks) and the functional impacts (e.g., mixes of land uses). This might be useful for land administration (including subdivisions) but, if strata title properties are to be visualized or the public and private ownership of future developments needs to be analyzed by stakeholders, today's tools must be improved.

The ability to measure the capacities and implications of underground infrastructure (Qiao et al. 2019) and the above-ground services needed for huge future developments can be added through combinations of 3D cadastre data and BIM methods. Adopting a 3D cadastre (Aien et al. 2013) will enable users to more accurately evaluate future changes in land and property values, which is a major concern for many stakeholders involved in inner city redevelopment (Shin 2009).

16.5.2 Graph Databases: Lossless Processes for Data Cities

Graph databases have nodes, edges and properties to represent and store data. They aid in analysis of many-many relationships and have applications in machine learning, fraud detection, social media, semantic harmonization and master data management. Graphs are a key enabler of next-generation spatial platforms for the integrated digital built environment (IDBE).

Recent Singapore research demonstrates potential for multidirectional lossless transformation of semantic and geometric data across the paradigms from design models to open standard formats. With traditional methods, there may be significant loss of data integrity and content at each step in the transformation from the proprietary design files (native BIM) into various OpenBIM standards (IFCs), and then into CityGML and a city model. This journey wrangles data between multiple paradigms. By adopting a triple graph-based framework for semantic and geometric conversions a "complete and near-lossless" mapping between the models can be achieved (Stouffs 2018). This framework can be applied to bulk and incremental updates between these models and may be applicable to lossless transformations between IFC versions (for example, IFC2x3 to IFC4x2) and pivoting from project information models (PIMs) to asset information models (AIMs) at handover.

Also significant for cities and major infrastructure projects and operations are graph application platforms for master data management (MDM). Property type graphs (with metadata at nodes and edges) will also become increasingly important for building asset registers and to enable sophisticated twinning of a sentient virtual model with its physical counterpart in the real world.

16.5.3 Open Public Life Data Protocol

Denmark's Gehl Institute, founded by Jan Gehl and now led by Shin-pei Tsay, has collaborated with city government agencies in Copenhagen, San Francisco and Seattle to produce a data protocol for assembling and comparing metrics about how people use and enjoy public environments. Launched as a beta version in late 2017, it includes a choice of eight components relevant to any public life survey: gender, age, mode of mobility, groups, posture, activities, objects (accompanying people) and "geotag" (which parts of a location are preferred). The protocol document explained how to structure the survey to record all the information in data tables, to be saved as CSV files that could be compared with the results of any similarly assembled survey. Although this project was focused on using digital tools to collect and process the survey information, it appears to be strongly influenced by Christopher Alexander's pattern language classifications (published before personal computers) to help design comfortable indoor and outdoor places at different scales and for different times and purposes (Gehl 2017; Alexander et al. 1977).

16.5.4 City Standards: ISO 37120

In 2014, the International Organization for Standardization (ISO) launched its first suite of indicators to measure and compare the performance of cities across seventeen general themes including the economy, energy, governance, health, telecommunications and innovation, transport, waste and water. Developed by the Global City Indicators Facility (GCIF) at the University of Toronto and promoted by the related World Council on City Data (WCCD), the ISO 37120 standard for sustainable cities and communities was updated in 2018. Two subset standards documents, *ISO 37122: Indicators for Smart Cities* and *ISO 37123: Indicators for Resilient Cities*, were also produced (ISO 2018).

16.5.5 Data Cubes

Launched in 2013 by the national EO team at Geoscience Australia, the Australian Geoscience Space-Time Data Cube is a system that stacks matching Landsat scenes in time sequences (currently up to fifteen years) to allow for faster analysis of changing conditions. The dataset for the whole of Australia amounts to almost four million scenes and 110 TB of compressed geoTIFF files, which are analyzed by the Raijin high-performance computing lab in Canberra. Technicians can access the dataset with a Python API that can generate specific mosaics and stacks of image files, which can be interpreted via users' own algorithms (Jackson 2013). The Data Cube system is the foundation of the Digital Earth Australia national satellite mapping project. It is supplied freely to research agencies in other countries under the name Open Data Cube (ODC), under the auspices of the Committee for Earth Observation Satellites (CEOS n.d.). One ODC was repurposed as the Africa Regional Data Cube, providing satellite surveys to an initial group of five African nations; it was later expanded to all 54 African countries as the Digital Earth Africa project (Digital Earth Africa n.d.)

In 2018, Peter Baumann at Jacobs University in Bremen received German research funding to lead a BigDataCube project to improve Rasdaman (raster data manager) software for data cubing satellite imagery from the European Space Agency's Sentinel constellation (source of the six terabytes per day of new satellite image files that are stored in Germany's CODE-DE archive). In addition to the commercial and free/light versions of Rasdaman, Baumann is developing data cube standards for the Open Geospatial Consortium (OGC), which works with the International Standards Organization (ISO; Anon 2018).

16.5.6 Compact Cities

In 2012, the Organization for Economic Cooperation and Development (OECD) released a list of fourteen characteristics of compact cities, which could be used as indicators to compare and improve operations. Compact and often high-rise cities such as Manhattan, Hong Kong and Paris are more efficient than sprawling low-rise cities with wide traffic thoroughfares, such as Los Angeles and Dubai. The criteria are high residential and employment densities, mixtures of land uses, a fine grain of land uses (small sizes of land parcels), strong social and economic interaction, contiguous development (rather than vacant land or street-level carparks), contained urban development within demarcated limits, good urban infrastructure (especially sewage and water mains), multimodal transport, high connectivity of streets (including footpaths and bicycle lanes), extensive coverage of impervious surfaces and a low ratio of open space (OECD 2012).

16.5.7 EcoDesign

Malaysian architect Ken Yeang pioneered ecological strategies for commercial towers and urban precincts. He clarified an "endemic" (climate and place-responsive) design system that is now standard practice in urban development. He rejected key modernist routines for tall buildings to look the same on all four sides, be built around a central lift core, have sealed windows and be mechanically air conditioned throughout. Instead, he designed buildings to respond to their different compass aspects and sun and wind conditions; positioned lift cores on the sides of buildings where they could best block excessive sun and wind and would allow for courtyards or atria to be landscaped in the center; introduced natural sunlight and ventilation via openable windows to the foyer, lift lobbies, fire stairs and toilets; and designed sky gardens and sunny courtyards on upper levels. By working closely with engineers on climate-response tests of his building models, he discovered that sky gardens could break the flows of winds down the surfaces of his towers to reduce gusts for pedestrians walking on nearby streets and plazas. On upper floor levels, some winds could be deflected into the buildings to ventilate spaces and cool the structures.

Yeang first applied green architecture principles to his tall buildings in tropical cities—beginning with Menara Mesiniaga, an office building in Kuala Lumpur (1992). His most substantial book, *EcoDesign* (Yeang 2006), was the first to clarify an eco-scientific system to design site-sensitive architectural projects anywhere in the world. His method requires understanding the natural context of each site by identifying its biome (regional community of diverse species sharing the same climate and terrestrial conditions). His system also included an “interactions matrix” that requires four sets of data to be gathered to assess four main ecological impacts of a building scheme: its relations to its environment, its internal relations, its inputs (of energy and matter) and its outputs. He urged architects to plan developments to avoid destroying healthy ecosystems or to rehabilitate damaged ecosystems. He also described three criteria for modeling any design: a description of the built system, a description of its environment and a mapping of interactions between the building and its environment (Yeang 2006, 59–73).

16.5.8 Positive Development

Counterproductive practices in the “sustainable urban development” movement have been targeted by Janis Birkeland, author of an evidence-based ecological building theory that she named “positive development” and “net-positive design” (Birkeland 2008). She proposed that every building should be expected to sequester the amount of carbon used in its operations and the amount of carbon emitted through resource extraction and consumption. Every building project destroys many tons of the Earth’s natural resources and the link between mining and construction is the major global cause of excessive carbon emissions. The nature and extent of this problem are obscured by green building assessment practices that “measure the wrong things in the wrong ways”—and that measure negative impacts only up to zero without measuring positive impacts. Birkeland suggested that buildings that support substantial and permanent planting (green walls and roofs) will amortize carbon far earlier in their life cycles than if they are only operated with renewable energy sources. She expected machine-analyzed data to allow for much more comprehensive and accurate analyses of buildings before and after construction, but the issues being recorded and assessed must be changed and expanded.

16.5.9 Cities and the Digital Earth Nervous System

In a 2010 article for the *International Journal of Digital Earth*, European scientists explained Digital Earth as a “metaphor for the organisation and access to digital information through a multiscale, three-dimensional representation of the globe” (De Longueville et al. 2010). They extended that vision by forecasting a “self-aware nervous system” to provide decision makers with improved alerting mechanisms for crisis prediction and situational awareness. This goal may prove the most beneficial for governments and citizens in urban areas—and strategies are essential to clarify how the relevant officials and stakeholders can contribute to and usefully exploit such a sophisticated system of automated operations.

New international protocols are essential to give urban authorities rapid access and accurate, automatic analyses of the data sets relevant to their challenges. Several international corporations—including Mapbox, Orbital Insight and OmniSci—are managing and analyzing very large quantities of geodata for government customers which cannot afford or do not find it feasible to apply the necessary resources and infrastructure to maintain such sophisticated operations.

One geospatially advanced city (and nation) is Singapore, which ranked fourth in the 2018 countries geospatial readiness index—not far behind the United States, United Kingdom and Germany, and ahead of China (Geospatial Media and Communications 2018). Its extraordinarily integrated government created a national spatial data infrastructure (NSDI) system in 2009, is locally training urban geotech specialists through joint research programs with Switzerland's ETH Future Cities Lab and America's MIT SENSEable Cities Lab (with centers located at the National University of Singapore), has accelerated a Smart Nation policy since 2014, launched the Virtual Singapore project in 2016 and released the Singapore Geospatial Masterplan in 2018. Singapore aims to foster "geosmart government", "geoempowered people", and "a thriving geoindustry".

16.6 Summary

International scientists supporting the Digital Earth and GEOSS visions are applying satellite and semiconductor-enabled technology to accelerate delivery of Fuller's visions for a "4D Air-Ocean World Town Plan" and efficient management of resources on "Spaceship Earth". Many governments have been promoting "smart city" policies and programs since the 1990s, when wireless and mobile telecommunications began to be commercialized internationally. The authors of this paper suggest it is now important to not only emphasize systems that enable humans to communicate worldwide but also next-generation infrastructure for societies to be accurately informed about our planet's environmental conditions and challenges.

At the urban scale of today's planet-simulation project, there is a need to integrate area-specific modeling of natural environmental systems with current best practices in building information modeling and city information modeling. All three methods must be improved to incorporate real-time streaming of Earth observations data obtained via sensing and scanning the light and radio waves of the electromagnetic spectrum. This satellite and semiconductor-enabled movement has been labeled "the new science of cities", "geodesign", "senseable cities", "digital cities", and "data cities". As De Longueville et al. clarified in 2010, it seems crucial for the Digital Earth "nervous system" to become self-aware and be able to obtain and respond more automatically to unprecedented quantities of environmental information—far too much to be processed by humans. This chapter highlighted some urban advances, strategies, issues and case studies that are significant contributions to this millennium's Digital Earth/GEOSS imperative.

References

Adriaenssens S, Gramazio F, Kohler M et al (2016) Advances in architectural geometry 2016. https://vdf.ch/advances-in-architectural-geometry-2016-e-book.html. Accessed 23 Jan 2019

Agius T, Sabri S, Kalantari M (2018) Three-dimensional rule-based city modelling to support urban redevelopment process. ISPRS Int J Geo-Inf 7(10):413

Aiello LM, Schifanella R, Quercia D et al (2016) Chatty maps: constructing sound maps of urban areas from social media data. R Soc Open Sci 3(3):150690

Aien A, Kalantari M, Rajabifard A et al (2013) Utilising data modelling to understand the structure of 3D cadastres. J Spat Sci 58(2):215–234

Alexander C, Ishikawa S, Silverstein M (1977) A pattern language: towns buildings construction. Oxford University Press, New York

Anon (2013) Artificial sun brightens up Arctic, 27 September, 47, Metro (London)

Anon (2018) Observing the earth via satellite: cubes to sort data, Jacobs university. https://www.jacobs-university.de/news/observing-earth-satellite-cubes-sort-data. Accessed 29 Jan 2019

Batty M (2005) Cities and complexity: understanding cities with cellular automata, agent-based models, and fractals. The MIT Press, Cambridge, MA

Batty M (2013) The new science of cities. The MIT Press, Cambridge, MA

Batty M (2018) Artificial intelligence and smart cities. Environ Plan B Urban Anal City Sci 45(1):3–6

Batty M, Longley P (1994) Fractal cities: a geometry of form and function. Academic Press, San Diego, CA

Betsky A, Aadigard E (2000) Architecture must burn: a manifesto for architecture beyond building. Thames and Hudson, London

Biljecki F, Ledoux H, Stoter J et al (2014) Formalisation of the level of detail in 3D city modelling. Comput Environ Urban Syst 48:1–15

Birkeland J (2008) Positive development: from vicious circles to virtuous cycles through built environment design. Earthscan, London

Bishop R (2018) Earth news from the press at large, ICES foundation. http://www.icesfoundation.org/Pages/Home.aspx. Accessed 23 Apr 2018

Borrmann D, Nüchter A (n.d.) No 22, robotic 3D scan repository. Accessed 29 Jan 2019

Brosan R, Segal G (1982) Machine of the year: the computer moves in, Time. http://content.time.com/time/covers/0,16641,19830103,00.html. Accessed 11 Jul 2019

Committee for Earth Observation Satellites (CEOS) (n.d.) Open data cube. https://www.opendatacube.org/about. Accessed 30 Jan 2019

Craglia M, de Bie K, Jackson D et al (2012) Digital earth 2020: towards the vision for the next decade. Int J Digit Earth 5(1):4–21

Crompvoets J, Rajabifard A, Loenen BV et al (2008) A multi-view framework to assess spatial data infrastructures. Waginengen University Space for Geo-Information, Waginengen (Belgium) and University of Melbourne Department of Geomatics, Melbourne

Dangermond J (2018) Vision: applying the science of where to geodesign. http://proceedings.esri.com/library/userconf/geodesign18/papers/geodesign-01.pdf. Accessed 23 Jan 2019

Daniels M (2018a) Human terrain: visualizing the world's population in 3D, The pudding. https://pudding.cool/2018/10/city_3d/. Accessed 21 Jan 2019

Daniels M (2018b) 3D mapping global population density: how i built it, Mapbox. https://blog.mapbox.com/3d-mapping-global-population-density-how-i-built-it-141785c91107. Accessed 21 Jan 2019

De Longueville B, Annoni A, Schade S et al (2010) Digital earth's nervous system for crisis events: real-time sensor web enablement of volunteered geographic information. Int J Digit Earth 3(3):242–259

De Monchaux N (2016) Local code: 3,659 proposals about data, design and the nature of cities. Princeton Architectural Press, New York

Digital Earth Africa (n.d.) Digital earth Africa, Geoscience Australia. http://www.ga.gov.au/digitalearthafrica. Accessed 30 Jan 2019

DLR (German Aerospace Center) (n.d.) Global urban footprint, DLR earth observations center. https://www.dlr.de/eoc/en/desktopdefault.aspx/tabid-9628/16557_read-40454/. Accessed 21 Jun 2019

Droege P (1997) Intelligent environments: spatial aspects of the information revolution. North-Holland/Elsevier Science, Amsterdam

Forrester JW (1969) Urban dynamics. The MIT Press, Cambridge, MA

Fuller RB (1928) 4D timelock. Privately Published, Chicago

Fuller RB (1969) Operating manual for spaceship earth. Southern Illinois University Press, Carbondale, IL

Fuller RB (1980) Critical path. St Martin's Press, New York, NY

Gates W (1995) The road ahead. Viking, New York

Gehl Institute (2017) The open public life data protocol. Gehl Institute, Copenhagen

Geospatial Media and Communications (2018) Countries geospatial readiness index, geo-buiz geospatial industry outlook and readiness index. GMC, Amsterdam

Gibson DV, Kozmetsky G, Smilor RW (1992) The technopolis phenomenon: smart cities, fast systems, global networks. Rowman and Littlefield, Lanham, MD

Goodchild MF, Guo H, Annoni A et al (2012) Next-generation digital earth. Proc Natl Acad Sci U S A 109(28):11088–11094

Gore A (1992) Earth in the balance: ecology and the human spirit. Houghton Mifflin, New York, NY

Gore A (1999) The digital earth: understanding our planet in the 21st century, photogrammetric engineering & remote sensing. https://www.asprs.org/wp-content/uploads/pers/99journal/may/1999_may_highlight.pdf. Accessed 23 Jan 2019

Group on Earth Observations (2007) The full picture. GEO, Geneva

Group on Earth Observations (2015) GEO strategic plan 2016–2025: implementing GEOSS. GEO, Geneva

Haas H (2011) Wireless data from every light bulb, TED: ideas worth spreading. https://www.ted.com/talks/harald_haas_wireless_data_from_every_light_bulb/footnotes?awesm=on.ted.com_HHaas&utm_campaign=&utm_medium=on.ted.com-static&utm_source=direct-on.ted.com&utm_content=awesm-bookmarklet. Accessed 19 Mar 2018

Haas H (2015) Forget Wi-Fi. Meet the new Li-Fi internet, TED: ideas worth spreading. https://www.ted.com/talks/harald_haas_a_breakthrough_new_kind_of_wireless_internet. Accessed 19 Mar 2018

Hunter F, Hartcher P (2019) Revealed: Australia's key role in China's rival GPS. The Sydney Morning Herald, Pyrmont, New South Wales

International Organization for Standardization (ISO) (2018) ISO 37120:2018. Sustainable cities and communities—indicators for city services and quality of life. https://www.iso.org/standard/68498.html. Accessed 30 Jan 2019

Ito K, Ota H, Okabe T et al (2005a) Visualization of an urbanizing and globalizing world with the cityscape metaphor. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.623.8210&rep=rep1&type=pdf. Accessed 21 Jan 2019

Jackson D (2008) D_City: networking the data modelling revolution. In: Johnson C, Hu R, Abedin S, (eds) Connecting cities: networking Metropolis congress 2008 publication. NSW Government, Sydney, p 154–175, https://www.metropolis.org/sites/default/files/network_complete_0.pdf. Accessed 3 August 2019

Jackson D (2013) Data cube: GA crunches landsat tiles in time stacks virtual ANZ. http://virtualanz.net/data-cube-ga-crunches-landsat-tiles-in-time-stacks/. Accessed 29 Jan 2019
Jackson D (ed) (2015) SuperLux: smart light art, design and architecture for cities. Thames and Hudson, London
Jackson D (2016) Rebooting spaceship earth: astrospatial visions for architecture and urban design. In: Cairns G, (ed) Visioning technologies: the architectures of sight. Routledge, Abingdon, OX, p 121–135
Jackson D (2018) Data cities: how satellites are transforming architecture and design. Lund Humphries, London
Jackson D, Simpson R (2012) D_City: digital earth|virtual nations|data cities: global futures for environmental planning. http://www.dcitynetwork.net/manifesto. Accessed 18 Jan 2019
Kohler M, Gramazio F, Willmann J (2014) The robotic touch: how robots change architecture. Park Books, Zurich
Kurzweil R (2005) The singularity is near: when humans transcend biology. Viking, New York, NY
Lin Y-T, Maire M, Belongie S, Bourdev L, Girshick R, Hays J, Perona P, Ramanan D, Zitnick CL, Dollár P (2015) Microsoft COCO: Common Objects in Context https://arxiv.org/pdf/1405.0312.pdf. Accessed 2 August 2019
MIT SENSEable City Lab (2008) NYTE: New York talk exchange, Senseable City Lab: MIT. http://senseable.mit.edu/nyte/. Accessed 21 Jan 2019
Mitchell WJ (1995) City of bits: space, place and the infobahn. The MIT Press, Cambridge, MA
Modha DS (n.d.) Introducing a brain-inspired computer, IBM. http://www.research.ibm.com/articles/brain-chip.shtml?utm_content=buffer2756c&utm_medium=social&utm_source=twitter.com&utm_campaign=buffer. Accessed 24 Apr 2018
MVRDV (1999) Metacity datatown. 010, Rotterdam
MVRDV (n.d.) Almere 2030. https://www.mvrdv.nl/projects/357/almere-2030. Accessed 29 Jan 2019
MVRDV, Delft School of Design, Berlage Institute et al (2007) Spacefighter: the evolutionary city (game). Actar, Barcelona
Negroponte N (1993) HDTV: what's wrong with this picture? Wired. https://web.media.mit.edu/~nicholas/Wired/WIRED1-01.html. Accessed 18 Jan 2019
Negroponte N (1995) Being digital: the roadmap for survival on the information superhighway. Hodder and Stoughton, New York
Neumann D, Champa KS (2002) Architecture of the night: the illuminated building. Prestel, New York
Organization for Economic Cooperation and Development (OECD) (2012) Compact city policies: a comparative assessment. OECD, Paris
Paradiso J (2011) Doppellab: spatialized sonification in a 3D virtual environment. https://www.media.mit.edu/projects/doppellab-spatialized-sonification-in-a-3d-virtual-environment/overview/. Accessed 12 Jul 2019
Pesaresi M, Ehrlich D, Ferri S et al (2016) Operating procedure for the production of the global human settlement layer from landsat data of the epochs 1975, 1990, 2000 and 2014 (EC JRC technical report). European Commission Joint Research Center, Ispra
Peters B, Peters T (2013) Inside smartgeometry: expanding the architectural possibilities of computational design: AD Smart01. John Wiley and Sons, London
Qiao Y-K, Peng F-L, Sabri S et al (2019) Low carbon effects of urban underground space. Sustain Cities Soc 45:451–459
Quercia D, Aiello LM, Schifanella R (2016) The emotional and chromatic layers of urban smells. In: Proceedings of the 10th international AAAI conference on web and social media (ICWSM 2016), North America, pp 309–318
Rae A (2016) The global human settlement layer: an amazing new global population dataset, Stats, Maps n Pix. http://www.statsmapsnpix.com/2016/. Accessed 21 Jan 2019
Rae A (2018) One degree of population, Stats, Maps n Pix. http://www.statsmapsnpix.com/. Accessed 21 Jan 2019

Rahm P (n.d.) Jade Eco park, Philippe Rahm architectes. http://www.philipperahm.com/data/projects/taiwan/index.html. Accessed 29 Jan 2019
Ratti C, Claudel M (2016) The city of tomorrow: sensors, networks, hackers, and the future of urban life. Yale University Press, New Haven, CT
Sabri S, Pettit CJ, Kalantari M et al (2015) What are essential requirements in planning for future cities using open data infrastructures and 3D data models? In: 14th computers in urban planning and urban management (CUPUM2015). MIT, Boston, MA, pp 314.1–314.17
Schumacher P (2008) Parametricism as style—parametricist manifesto. https://www.patrikschumacher.com/Texts/Parametricism%20as%20Style.htm. Accessed 23 Jan 2019
Shin HB (2009) Property-based redevelopment and gentrification: the case of Seoul, South Korea. Geoforum 40(5):906–917
Smartgeometry (n.d.) About. https://www.smartgeometry.org/about. Accessed 12 Jul 2019
Stanford Vision Lab (2016) ImageNet. http://www.image-net.org/. Accessed 19 Mar 2018
Stanley A (2017) Virtual reality: The new reality for real estate, the urban developer. https://theurbandeveloper.com/articles/virtual-reality-the-new-reality. Accessed 4 Mar 2018
Statt N (2018) Vuzix Blade AR glasses are the next-gen google glass we've all been waiting for, the verge. https://www.theverge.com/2018/1/9/16869174/vuzix-blade-ar-glasses-augmented-reality-amazon-alexa-ai-ces-2018. Accessed 6 Mar 2018
Steinitz C (2012) A framework for geodesign: changing geography by design. Esri, Redlands, CA
Steinitz C (2013) Beginnings of geodesign: a personal historical perspective. In: Geodesign: past, present and future. Redlands, CA, Esri, pp 4–14
Stouffs R (2018) A triple graph grammar approach to mapping IFC models into CityGML building models. In: 23rd conference on computer-aided architectural design research in Asia (CAADRIA). Tsinghua University, Beijing, pp 41–50
Team PopulouSCAPE (2005) PopulouSCAPE: night flight over an urbanizing world. https://vimeo.com/populouscape. Accessed 21 Jan 2019
United Nations Development Program (UNDP) (n.d.) Our world is getting warmer. https://sdgs.undp.org/?utm_source=web&utm_medium=sdgs&utm_campaign=sdgstoday. Accessed 21 Jan 2019
United Nations Global Marketplace (UNGM) (n.d.) The UN system organizations, united nations global marketplace. https://www.ungm.org/Shared/KnowledgeCenter/Pages/VBS_UNSystem. Accessed 21 Jan 2019
Venkataramanan M (2014) Space archaeologist dvers lost cities with satellite imagery, Wired. https://www.wired.co.uk/article/scanning-the-past. Accessed 23 Jan 2019
Wolfram S (2002) A new kind of science. Wolfram Media, Champaign, Ill
Wyss M (n.d.) Near real-time earthquake loss estimates, ICES foundation. http://www.icesfoundation.org/Pages/QlarmEventList.aspx. Accessed 23 Apr 2018
Yeang K (2006) Ecodesign: a manual for ecological design. Wiley-Academy, London

Davina Jackson PhD is a Sydney-based writer on urban technology applications, an honorary academic with the University of Kent School of Architecture and a Fellow of the Royal Geographical Society. She co-edited the world's first snapshot report of academic contributions to the Global Earth Observation System of Systems/Digital Earth project and her latest book is Data Cities: How satellites are transforming architecture and design (Lund Humphries, 2018).

Richard Simpson is an Executive Director of Meta Moto, a consultancy firm specialized in advising governments on strategic spatial programs. He is a Senior Fellow at the Centre for Spatial Data Infrastructure and Land Administration (CSDILA) at the University of Melbourne, and at the Cities Research Institute, Griffith University; and expert advisor to the OGC-bSI joint working group for the Integrated Digital Built Environment. He has served on the ISDE Council since it was founded in 2006.

The Way Forward

We have an unparalleled opportunity to turn a flood of raw data into understandable information about our society and out planet. This data will include not only high-resolution satellite imagery of the planet, digital maps, and economic, social, and demographic information. If we are successful, it will have broad societal and commercial benefits in areas such as education, decision-making for a sustainable future, land-use planning, agricultural, and crisis management.

The Digital Earth project could allow us to respond to manmade or natural disasters—or to collaborate on the long-term environmental challenges we face.

A Digital Earth could provide a mechanism for users to navigate and search for geospatial information—and for producers to publish it. The Digital Earth would be composed of both the "user interface"—a browsable, 3D version of the planet available at various levels of resolution, a rapidly growing universe of networked geospatial information, and the mechanisms for integrating and displaying information from multiple sources.

A comparison with the World Wide Web is constructive. [In fact, it might build on several key Web and Internet standards.] Like the Web, the Digital Earth would organically evolve over time, as technology improves and the information available expands. Rather than being maintained by a single organization, it would be composed of both publically available information and commercial products and services from thousands of different organizations. Just as interoperability was the key for the Web, the ability to discover and display data contained in different formats would be essential.

I believe that the way to spark the development of a Digital Earth is to sponsor a testbed, with participation from government, industry, and academia. This testbed would focus on a few applications, such as education and the environment, as well as the tough technical issues associated with interoperability, and policy issues such as privacy. As prototypes became available, it would also be possible to interact with the Digital Earth in multiple places around the country with access to high-speed networks, and get a more limited level of access over the Internet.

Clearly, the Digital Earth will not happen overnight.

In the first stage, we should focus on integrating the data from multiple sources that we already have. We should also connect our leading children's museums and science museums to high-speed networks such as the Next Generation Internet so that children can explore our planet. University researchers would be encouraged to partner with local schools and museums to enrich the Digital Earth project—possibly by concentrating on local geospatial information.

Next, we should endeavor to develop a digital map of the world at 1 meter resolution.

In the long run, we should seek to put the full range of data about our planet and our history at our fingertips.

In the months ahead, I intend to challenge experts in government, industry, academia, and non-profit organizations to help develop a strategy for realizing this vision. Working together, we can help solve many of the most pressing problems facing our society, inspiring our children to learn more about the world around them, and accelerate the growth of a multi-billion dollar industry.

Appendix F
1999 Beijing Declaration on Digital Earth and 2009 Beijing Declaration on Digital Earth

Beijing Declaration on Digital Earth

December 2, 1999

We, some 500 scientists, engineers, educators, managers and industrial entrepreneurs from 20 countries and regions assembled here in the historical city of Beijing, attending the first International Symposium on Digital Earth being organized by the Chinese Academy of Sciences with co-sponsorship of 19 organizations and institutions from November 29, 1999 to December 2, 1999, recognize that humankind, while entering into the new millennium, still faces great challenges such as rapid population growth, environmental degradation, and natural resource depletion which continue to threaten global sustainable development;

Noting that global development in the 20th century has been characterized by rapid advancements in science and technology which have made significant contributions to economic growth and social wellbeing and that the new century will be an era of information and space technologies supporting the global knowledge economy;

Recalling the statement by Al Gore, Vice President of the United States of America, on *Digital Earth: Understanding Our Planet in the 21st Century*—and the statement by Jiang Zemin, President of the People's Republic of China, on Digital Earth regarding trends of social, economic, scientific and technological development;

Realizing the decisions made at UNCED and Agenda 21, recommendations made by UNISPACE III and the Vienna Declaration on Space and Human Development, which address, among other things, the importance of the Integrated Global Observing Strategy, the Global Spatial Data Infrastructure, geographic information systems, global navigation and positioning systems, geo-spatial information infrastructures and modeling of dynamic processes;

Understanding that Digital Earth, addressing the social, economic, cultural, institutional, scientific, educational, and technical challenges, allows humankind to visualize the Earth, and all places within it, to access information about it and to understand and influence the social, economic and environmental issues that affect their lives in their neighborhoods, their nations and the planet Earth;

H. Guo et al. (eds.), *Manual of Digital Earth*,
https://doi.org/10.1007/978-981-32-9915-3

Recommend that Digital Earth be promoted by scientific, educational and technological communities, industry, governments, as well as regional and international organizations;

Recommend also that while implementing the Digital Earth, priority be given to solving problems in environmental protection, disaster management, natural resource conservation, and sustainable economic and social development as well as improving the quality of life of the humankind;

Recommend further that Digital Earth be created in a way that also contributes to the exploration of, and scientific research on, global issues and the Earth system;

Declare the importance of Digital Earth in achieving global sustainable development;

Call for adequate investments and strong support in scientific research and development, education and training, capacity building as well as information and technology infrastructures, with emphasis, inter alia, on global systematic observation and modeling, communication networks, database development, and issues associated with interoperability of geo-spatial data;

Further call for close cooperation and collaboration between governments, public and private sectors, non-governmental organizations, and international organizations and institutions, so as to ensure equity in distribution of benefits derived from the use of Digital Earth in developed and developing economies;

Agree that, as a follow-up to the first International Symposium on Digital Earth held in Beijing, the International Symposium on Digital Earth should continue to be organized by interested countries or organizations biannually, on a rotational basis.

Beijing Declaration on Digital Earth
September 12, 2009

We scientists, engineers, educators, entrepreneurs, managers, administrators and representatives of civil societies from more than forty countries, international organizations and NGOs, once again, have assembled here, in the historic city of Beijing, to attend the Sixth International Symposium on Digital Earth, organized by the International Society for Digital Earth and the Chinese Academy of Sciences, with co-sponsorship of sixteen Chinese Government Departments, Institutions and international organizations, being held from September 9–12, 2009.

Noting

That Significant global-scale developments on Digital Earth science and technology have been made over the past ten years, and parallel advances in space information technology, communication network technology, high-performance computing, and Earth System Science have resulted in the rise of a Digital Earth data-sharing platform for public and commercial purposes, so that now Digital Earth is accessible by hundreds of millions, thus changing both the production and lifestyle of mankind;

Recognizing

The contributions to Digital Earth made by the host countries of the previous International Symposia on Digital Earth since November 1999, including China, Canada, the Czech Republic, Japan and the USA, and by the host countries of the previous Summit Conferences on Digital Earth, including New Zealand and Germany, for the success of the meetings as well as further promotion of Digital Earth;

Further, that the establishment of the International Society for Digital Earth and the accomplishments of its Executive Committee, the launch of the International Journal on Digital Earth, and its global contribution to cooperation and data exchange;

That the themes of the previous seven meetings: Moving towards Digital Earth, Beyond Information Infrastructure, Information Resources for Global Sustainability, Digital Earth as Global Commons, Bring Digital Earth down to Earth, Digital Earth and Sustainability, Digital Earth and Global Change, and Digital Earth in Action, have laid out a panoramic scenario for the future growth of Digital Earth;

That Digital Earth will be asked to bear increased responsibilities in the years to come, in the face of the problems of sustainable development;

Further Recognizing

That Digital Earth should play a strategic and sustainable role in addressing such challenges to human society as natural resource depletion, food and water insecurity, energy shortages, environmental degradation, natural disasters response, population explosion, and, in particular, global climate change;

That the purpose and mission of the World Information Summit of 2007, the Global Earth Observation System Conference of 2007, and the upcoming United Nations Climate Change Conference of 2009, and that Digital Earth is committed to continued close cooperation with other scientific disciplines;

Realizing

That Digital Earth is an integral part of other advanced technologies including: earth observation, geo-information systems, global positioning systems, communication networks, sensor webs, electromagnetic identifiers, virtual reality, grid computation, etc. It is seen as a global strategic contributor to scientific and technological developments, and will be a catalyst in finding solutions to international scientific and societal issues;

We Recommend

(a) That Digital Earth expand its role in accelerating information transfer from theoretical discussions to applications using the emerging spatial data infrastructures worldwide, in particular, in all fields related to global climate change, natural disaster prevention and response, new energy-source development, agricultural and food security, and urban planning and management;
(b) Further, that every effort be undertaken to increase the capacity for information resource-sharing and the transformation of raw data to practical information and applications, and developed and developing countries accelerate their programs to assist less-developed countries to enable them to close the digital gap and enable information sharing;

(c) Also, that in constructing the Digital Earth system, efforts must be made to take full advantage of next-generation technologies, including: earth observation, networking, database searching, navigation, and cloud computing to increase service to the public and decrease costs;
(d) Further, that the International Society for Digital Earth periodically take the lead in coordinating global scientific research, consultations and popular science promotion to promote the development of Digital Earth;
(e) Expanding cooperation and collaboration between the International Society for Digital Earth and the international community, in particular with intergovernmental organizations, and international non-governmental organizations;
(f) Extending cooperation and integration with Government Departments, the international Scientific and Educational community, businesses and companies engaged in the establishment of Digital Earth;

We Call for

Support from planners and decision-makers at all levels in developing plans, policies, regulations, standards and criteria related to Digital Earth, and appropriate investments in scientific research, technology development, education, and popular promotion of the benefits of Digital Earth.